PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

B3

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C07K 1/00, C07H 21/02, A61K 39/00, C12N 1/20

A1

(11) International Publication Number: **WO 98/57976**

(43) International Publication Date: 23 December 1998 (23.12.98)

(21) International Application Number: PCT/US98/12516

(22) International Filing Date: 18 June 1998 (18.06.98)

(30) Priority Data:

08/878,715	19 June 1997 (19.06.97)	US
09/098,588	17 June 1998 (17.06.98)	US

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(71) Applicant: GENETICS INSTITUTE, INC. [US/US]; 87 CambridgePark Drive, Cambridge, MA 02140 (US).

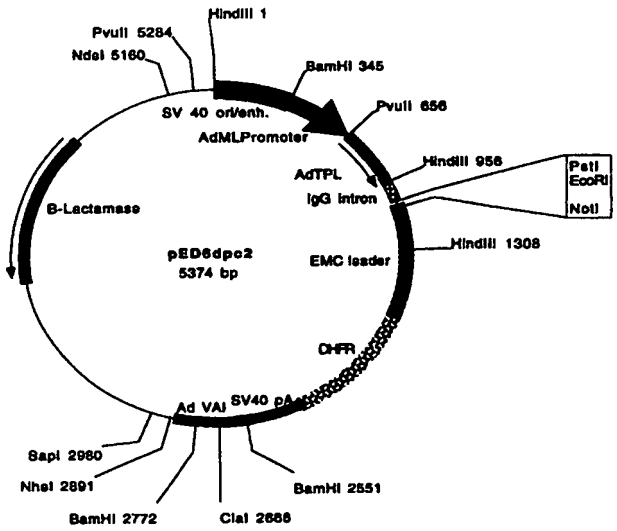
(72) Inventors: JACOBS, Kenneth; 151 Beaumont Avenue, Newton, MA 02160 (US). MCCOY, John, M.; 56 Howard Street, Reading, MA 01867 (US). LAVALLIE, Edward, R.; 113 Ann Lee Road, Harvard, MA 01451 (US). RACIE, Lisa, A.; 124 School Street, Acton, MA 01720 (US). TREACY, Maurice; 93 Walcott Road, Chestnut Hill, MA 02167 (US). SPAULDING, Vikki; 11 Meadowbank Road, Billerica, MA 01821 (US). AGOSTINO, Michael, J.; 26 Wolcott Avenue, Andover, MA 01810 (US). HOWES, Steven, H.; Apartment 2, 44 Chester Street, Somerville, MA 02144 (US). FECHTEL, Kim; 46 Marion Road, Arlington, MA 02174 (US).

(74) Agent: SPRUNGER, Suzanne, A.; Genetics Institute, Inc., 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(54) Title: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(57) Abstract

Novel polynucleotides and the proteins encoded thereby are disclosed.



Plasmid name: pED6dpc2
Plasmid size: 5374 bp

Comments/References: pED6dpc2 is derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al.(1991), NAR 19: 4485-4490.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

This application is a continuation-in-part of application Ser. No. 60/XXX,XXX (converted to a provisional application from non-provisional application Ser. No. 08/878,715), filed June 19, 1997, which is incorporated by reference herein.

FIELD OF THE INVENTION

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins.

BACKGROUND OF THE INVENTION

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, CSFs and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity by virtue of their secreted nature in the case of leader sequence cloning, or by virtue of the cell or tissue source in the case of PCR-based techniques. It is to these proteins and the polynucleotides encoding them that the present invention is directed.

SUMMARY OF THE INVENTION

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 185 to nucleotide 1600;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 1403 to nucleotide 1600;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 1 to nucleotide 850;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone do15_4 deposited under accession number ATCC 98468;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone do15_4 deposited under accession number ATCC 98468;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone do15_4 deposited under accession number ATCC 98468;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone do15_4 deposited under accession number ATCC 98468;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:2;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:1 from nucleotide 185 to nucleotide 1600; the nucleotide sequence of SEQ ID NO:1

from nucleotide 1403 to nucleotide 1600; the nucleotide sequence of SEQ ID NO:1 from nucleotide 1 to nucleotide 850; the nucleotide sequence of the full-length protein coding sequence of clone do15_4 deposited under accession number ATCC 98468; or the nucleotide sequence of a mature protein coding sequence of clone do15_4 deposited under accession number ATCC 98468. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone do15_4 deposited under accession number ATCC 98468. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 222. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:2, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment comprising the amino acid sequence from amino acid 231 to amino acid 240 of SEQ ID NO:2.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:1.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:2;

(b) the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 222;

(c) fragments of the amino acid sequence of SEQ ID NO:2 comprising eight consecutive amino acids of SEQ ID NO:2; and

(d) the amino acid sequence encoded by the cDNA insert of clone do15_4 deposited under accession number ATCC 98468;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:2 or the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 222. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID

NO:2, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2having biological activity, the fragment comprising the amino acid sequence from amino acid 231 to amino acid 240 of SEQ ID NO:2.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 47 to nucleotide 2065;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 1086 to nucleotide 1848;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone dx290_1 deposited under accession number ATCC 98468;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone dx290_1 deposited under accession number ATCC 98468;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone dx290_1 deposited under accession number ATCC 98468;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone dx290_1 deposited under accession number ATCC 98468;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:4;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:3 from nucleotide 47 to nucleotide 2065; the nucleotide sequence of SEQ ID NO:3

from nucleotide 1086 to nucleotide 1848; the nucleotide sequence of the full-length protein coding sequence of clone dx290_1 deposited under accession number ATCC 98468; or the nucleotide sequence of a mature protein coding sequence of clone dx290_1 deposited under accession number ATCC 98468. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone dx290_1 deposited under accession number ATCC 98468. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4 from amino acid 312 to amino acid 600. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:4, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment comprising the amino acid sequence from amino acid 331 to amino acid 340 of SEQ ID NO:4.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:3.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:4;

(b) the amino acid sequence of SEQ ID NO:4 from amino acid 312 to amino acid 600;

(c) fragments of the amino acid sequence of SEQ ID NO:4 comprising eight consecutive amino acids of SEQ ID NO:4; and

(d) the amino acid sequence encoded by the cDNA insert of clone dx290_1 deposited under accession number ATCC 98468;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:4 or the amino acid sequence of SEQ ID NO:4 from amino acid 312 to amino acid 600. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:4, or a protein comprising a fragment of the amino acid sequence of

SEQ ID NO:4having biological activity, the fragment comprising the amino acid sequence from amino acid 331 to amino acid 340 of SEQ ID NO:4.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 107 to nucleotide 724;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 218 to nucleotide 724;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 536 to nucleotide 866;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ek390_4 deposited under accession number ATCC 98468;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ek390_4 deposited under accession number ATCC 98468;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone ek390_4 deposited under accession number ATCC 98468;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone ek390_4 deposited under accession number ATCC 98468;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:6;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:5 from nucleotide 107 to nucleotide 724; the nucleotide sequence of SEQ ID NO:5 from nucleotide 218 to nucleotide 724; the nucleotide sequence of SEQ ID NO:5 from nucleotide 536 to nucleotide 866; the nucleotide sequence of the full-length protein coding sequence of clone ek390_4 deposited under accession number ATCC 98468; or the nucleotide sequence of a mature protein coding sequence of clone ek390_4 deposited under accession number ATCC 98468. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone ek390_4 deposited under accession number ATCC 98468. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6 from amino acid 6 to amino acid 92. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:6, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment comprising the amino acid sequence from amino acid 97 to amino acid 106 of SEQ ID NO:6.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:5.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:6;

(b) the amino acid sequence of SEQ ID NO:6 from amino acid 6 to amino acid 92;

(c) fragments of the amino acid sequence of SEQ ID NO:6 comprising eight consecutive amino acids of SEQ ID NO:6; and

(d) the amino acid sequence encoded by the cDNA insert of clone ek390_4 deposited under accession number ATCC 98468;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:6 or the amino acid sequence of SEQ ID NO:6 from amino acid 6 to amino acid 92. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid

sequence of SEQ ID NO:6 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:6, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6having biological activity, the fragment comprising the amino acid sequence from amino acid 97 to amino acid 106 of SEQ ID NO:6.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 31 to nucleotide 1230;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 289 to nucleotide 1230;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 344 to nucleotide 1119;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone er471_7 deposited under accession number ATCC 98468;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone er471_7 deposited under accession number ATCC 98468;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone er471_7 deposited under accession number ATCC 98468;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone er471_7 deposited under accession number ATCC 98468;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:8;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:7 from nucleotide 31 to nucleotide 1230; the nucleotide sequence of SEQ ID NO:7 from nucleotide 289 to nucleotide 1230; the nucleotide sequence of SEQ ID NO:7 from nucleotide 344 to nucleotide 1119; the nucleotide sequence of the full-length protein coding sequence of clone er471_7 deposited under accession number ATCC 98468; or the nucleotide sequence of a mature protein coding sequence of clone er471_7 deposited under accession number ATCC 98468. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone er471_7 deposited under accession number ATCC 98468. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8 from amino acid 111 to amino acid 363. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:8, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment comprising the amino acid sequence from amino acid 195 to amino acid 204 of SEQ ID NO:8.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:7.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:8;

(b) the amino acid sequence of SEQ ID NO:8 from amino acid 111 to amino acid 363;

(c) fragments of the amino acid sequence of SEQ ID NO:8 comprising eight consecutive amino acids of SEQ ID NO:8; and

(d) the amino acid sequence encoded by the cDNA insert of clone er471_7 deposited under accession number ATCC 98468;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:8 or the amino acid sequence

of SEQ ID NO:8 from amino acid 111 to amino acid 363. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:8, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8having biological activity, the fragment comprising the amino acid sequence from amino acid 195 to amino acid 204 of SEQ ID NO:8.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 62 to nucleotide 322;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 571 to nucleotide 878;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone fs40_3 deposited under accession number ATCC 98468;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone fs40_3 deposited under accession number ATCC 98468;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone fs40_3 deposited under accession number ATCC 98468;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone fs40_3 deposited under accession number ATCC 98468;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:10;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:10;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:9 from nucleotide 62 to nucleotide 322; the nucleotide sequence of SEQ ID NO:9 from nucleotide 571 to nucleotide 878; the nucleotide sequence of the full-length protein coding sequence of clone fs40_3 deposited under accession number ATCC 98468; or the nucleotide sequence of a mature protein coding sequence of clone fs40_3 deposited under accession number ATCC 98468. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone fs40_3 deposited under accession number ATCC 98468. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:10, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity, the fragment comprising the amino acid sequence from amino acid 38 to amino acid 47 of SEQ ID NO:10.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:9.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:10;

(b) fragments of the amino acid sequence of SEQ ID NO:10 comprising eight consecutive amino acids of SEQ ID NO:10; and

(c) the amino acid sequence encoded by the cDNA insert of clone fs40_3 deposited under accession number ATCC 98468;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:10. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:10, or a protein comprising a fragment of the amino acid sequence of

SEQ ID NO:10 having biological activity, the fragment comprising the amino acid sequence from amino acid 38 to amino acid 47 of SEQ ID NO:10.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 43 to nucleotide 1671;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 112 to nucleotide 1671;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 224 to nucleotide 679;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ga63_6 deposited under accession number ATCC 98468;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ga63_6 deposited under accession number ATCC 98468;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone ga63_6 deposited under accession number ATCC 98468;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone ga63_6 deposited under accession number ATCC 98468;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:12;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:11 from nucleotide 43 to nucleotide 1671; the nucleotide sequence of SEQ ID NO:11 from nucleotide 112 to nucleotide 1671; the nucleotide sequence of SEQ ID NO:11 from nucleotide 224 to nucleotide 679; the nucleotide sequence of the full-length protein coding sequence of clone ga63_6 deposited under accession number ATCC 98468; or the nucleotide sequence of a mature protein coding sequence of clone ga63_6 deposited under accession number ATCC 98468. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone ga63_6 deposited under accession number ATCC 98468. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12 from amino acid 62 to amino acid 212. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:12, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12 having biological activity, the fragment comprising the amino acid sequence from amino acid 266 to amino acid 275 of SEQ ID NO:12.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:11.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:12;

(b) the amino acid sequence of SEQ ID NO:12 from amino acid 62 to amino acid 212;

(c) fragments of the amino acid sequence of SEQ ID NO:12 comprising eight consecutive amino acids of SEQ ID NO:12; and

(d) the amino acid sequence encoded by the cDNA insert of clone ga63_6 deposited under accession number ATCC 98468;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:12 or the amino acid sequence of SEQ ID NO:12 from amino acid 62 to amino acid 212. In further preferred embodiments, the present invention provides a protein comprising a fragment of the

amino acid sequence of SEQ ID NO:12 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:12, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12having biological activity, the fragment comprising the amino acid sequence from amino acid 266 to amino acid 275 of SEQ ID NO:12.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 17 to nucleotide 523;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 77 to nucleotide 523;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 1 to nucleotide 392;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone gm335_4 deposited under accession number ATCC 98468;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone gm335_4 deposited under accession number ATCC 98468;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone gm335_4 deposited under accession number ATCC 98468;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone gm335_4 deposited under accession number ATCC 98468;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:14;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:14;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:13 from nucleotide 17 to nucleotide 523; the nucleotide sequence of SEQ ID NO:13 from nucleotide 77 to nucleotide 523; the nucleotide sequence of SEQ ID NO:13 from nucleotide 1 to nucleotide 392; the nucleotide sequence of the full-length protein coding sequence of clone gm335_4 deposited under accession number ATCC 98468; or the nucleotide sequence of a mature protein coding sequence of clone gm335_4 deposited under accession number ATCC 98468. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone gm335_4 deposited under accession number ATCC 98468. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:14 from amino acid 1 to amino acid 125. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:14, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14 having biological activity, the fragment comprising the amino acid sequence from amino acid 79 to amino acid 88 of SEQ ID NO:14.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:13.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:14;

(b) the amino acid sequence of SEQ ID NO:14 from amino acid 1 to amino acid 125;

(c) fragments of the amino acid sequence of SEQ ID NO:14 comprising eight consecutive amino acids of SEQ ID NO:14; and

(d) the amino acid sequence encoded by the cDNA insert of clone gm335_4 deposited under accession number ATCC 98468;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:14 or the amino acid sequence

of SEQ ID NO:14 from amino acid 1 to amino acid 125. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:14, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14having biological activity, the fragment comprising the amino acid sequence from amino acid 79 to amino acid 88 of SEQ ID NO:14.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 2 to nucleotide 991;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 62 to nucleotide 991;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 2 to nucleotide 504;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone hy370_9 deposited under accession number ATCC 98468;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone hy370_9 deposited under accession number ATCC 98468;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone hy370_9 deposited under accession number ATCC 98468;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone hy370_9 deposited under accession number ATCC 98468;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:16;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:15 from nucleotide 2 to nucleotide 991; the nucleotide sequence of SEQ ID NO:15 from nucleotide 62 to nucleotide 991; the nucleotide sequence of SEQ ID NO:15 from nucleotide 2 to nucleotide 504; the nucleotide sequence of the full-length protein coding sequence of clone hy370_9 deposited under accession number ATCC 98468; or the nucleotide sequence of a mature protein coding sequence of clone hy370_9 deposited under accession number ATCC 98468. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone hy370_9 deposited under accession number ATCC 98468. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 167. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:16, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity, the fragment comprising the amino acid sequence from amino acid 160 to amino acid 169 of SEQ ID NO:16.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:15.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:16;

(b) the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 167;

(c) fragments of the amino acid sequence of SEQ ID NO:16 comprising eight consecutive amino acids of SEQ ID NO:16; and

(d) the amino acid sequence encoded by the cDNA insert of clone hy370_9 deposited under accession number ATCC 98468;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:16 or the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 167. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:16, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16having biological activity, the fragment comprising the amino acid sequence from amino acid 160 to amino acid 169 of SEQ ID NO:16.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 77 to nucleotide 616;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 164 to nucleotide 616;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 1 to nucleotide 415;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ie47_4 deposited under accession number ATCC 98468;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ie47_4 deposited under accession number ATCC 98468;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone ie47_4 deposited under accession number ATCC 98468;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone ie47_4 deposited under accession number ATCC 98468;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:18;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:18;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:17 from nucleotide 77 to nucleotide 616; the nucleotide sequence of SEQ ID NO:17 from nucleotide 164 to nucleotide 616; the nucleotide sequence of SEQ ID NO:17 from nucleotide 1 to nucleotide 415; the nucleotide sequence of the full-length protein coding sequence of clone ie47_4 deposited under accession number ATCC 98468; or the nucleotide sequence of a mature protein coding sequence of clone ie47_4 deposited under accession number ATCC 98468. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone ie47_4 deposited under accession number ATCC 98468. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:18 from amino acid 1 to amino acid 113. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:18, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity, the fragment comprising the amino acid sequence from amino acid 85 to amino acid 94 of SEQ ID NO:18.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:17.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:18;

(b) the amino acid sequence of SEQ ID NO:18 from amino acid 1 to amino acid 113;

(c) fragments of the amino acid sequence of SEQ ID NO:18 comprising eight consecutive amino acids of SEQ ID NO:18; and

(d) the amino acid sequence encoded by the cDNA insert of clone ie47_4 deposited under accession number ATCC 98468;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:18 or the amino acid sequence of SEQ ID NO:18 from amino acid 1 to amino acid 113. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:18, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18having biological activity, the fragment comprising the amino acid sequence from amino acid 85 to amino acid 94 of SEQ ID NO:18.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 564 to nucleotide 2813;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 705 to nucleotide 2813;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 793 to nucleotide 1628;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone s195_10 deposited under accession number ATCC 98468;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone s195_10 deposited under accession number ATCC 98468;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone s195_10 deposited under accession number ATCC 98468;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone s195_10 deposited under accession number ATCC 98468;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:20;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:20;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:19 from nucleotide 564 to nucleotide 2813; the nucleotide sequence of SEQ ID NO:19 from nucleotide 705 to nucleotide 2813; the nucleotide sequence of SEQ ID NO:19 from nucleotide 793 to nucleotide 1628; the nucleotide sequence of the full-length protein coding sequence of clone s195_10 deposited under accession number ATCC 98468; or the nucleotide sequence of a mature protein coding sequence of clone s195_10 deposited under accession number ATCC 98468. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone s195_10 deposited under accession number ATCC 98468. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:20 from amino acid 78 to amino acid 355. In further preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:20, or a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20 having biological activity, the fragment comprising the amino acid sequence from amino acid 370 to amino acid 379 of SEQ ID NO:20.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:19.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:20;

(b) the amino acid sequence of SEQ ID NO:20 from amino acid 78 to amino acid 355;

(c) fragments of the amino acid sequence of SEQ ID NO:20 comprising eight consecutive amino acids of SEQ ID NO:20; and

(d) the amino acid sequence encoded by the cDNA insert of clone s195_10 deposited under accession number ATCC 98468;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:20 or the amino acid sequence of SEQ ID NO:20 from amino acid 78 to amino acid 355. In further preferred embodiments, the present invention provides a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20 having biological activity, the fragment preferably comprising eight (more preferably twenty, most preferably thirty) consecutive amino acids of SEQ ID NO:20, or a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20having biological activity, the fragment comprising the amino acid sequence from amino acid 370 to amino acid 379 of SEQ ID NO:20.

In certain preferred embodiments, the polynucleotide is operably linked to an expression control sequence. The invention also provides a host cell, including bacterial, yeast, insect and mammalian cells, transformed with such polynucleotide compositions. Also provided by the present invention are organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein.

Processes are also provided for producing a protein, which comprise:

(a) growing a culture of the host cell transformed with such polynucleotide compositions in a suitable culture medium; and

(b) purifying the protein from the culture.

The protein produced according to such methods is also provided by the present invention.

Protein compositions of the present invention may further comprise a pharmaceutically acceptable carrier. Compositions comprising an antibody which specifically reacts with such protein are also provided by the present invention.

Methods are also provided for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of a composition comprising a protein of the present invention and a pharmaceutically acceptable carrier.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A and 1B are schematic representations of the pED6 and pNOTs vectors, respectively, used for deposit of clones disclosed herein.

DETAILED DESCRIPTION

ISOLATED PROTEINS AND POLYNUCLEOTIDES

Nucleotide and amino acid sequences, as presently determined, are reported below for each clone and protein disclosed in the present application. The nucleotide sequence of each clone can readily be determined by sequencing of the deposited clone in accordance with known methods. The predicted amino acid sequence (both full-length and mature forms) can then be determined from such nucleotide sequence. The amino acid sequence of the protein encoded by a particular clone can also be determined by expression of the clone in a suitable host cell, collecting the protein and determining its sequence. For each disclosed protein applicants have identified what they have determined to be the reading frame best identifiable with sequence information available at the time of filing.

As used herein a "secreted" protein is one which, when expressed in a suitable host cell, is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence. "Secreted" proteins include without limitation proteins secreted wholly (e.g., soluble proteins) or partially (e.g. , receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins which are transported across the membrane of the endoplasmic reticulum.

Clone "do15_4"

A polynucleotide of the present invention has been identified as clone "do15_4". do15_4 was isolated from a human adult testes cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. do15_4 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "do15_4 protein").

The nucleotide sequence of do15_4 as presently determined is reported in SEQ ID NO:1. What applicants presently believe to be the proper reading frame and the predicted

amino acid sequence of the do15_4 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:2. Amino acids 394 to 406 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 407, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone do15_4 should be approximately 1900 bp.

The nucleotide sequence disclosed herein for do15_4 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. do15_4 demonstrated at least some similarity with sequences identified as AA113909 (zm80f12.r1 Stratagene neuroepithelium (#937231) Homo sapiens cDNA clone 531983 5'), AA189888 (mu55h06.r1 Soares mouse lymph node NbMLN Mus musculus cDNA clone 643355 5'), and U52052 (Human S6 A-8 mRNA expressed in chromosome 6-suppressed melanoma cells). Based upon sequence similarity, do15_4 proteins and each similar protein or peptide may share at least some activity.

Clone "dx290_1"

A polynucleotide of the present invention has been identified as clone "dx290_1". dx290_1 was isolated from a human adult testes cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. dx290_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "dx290_1 protein").

The nucleotide sequence of dx290_1 as presently determined is reported in SEQ ID NO:3. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the dx290_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:4.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone dx290_1 should be approximately 2300 bp.

The nucleotide sequence disclosed herein for dx290_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. dx290_1 demonstrated at least some similarity with the sequence identified as AA064383 (ml47h02.r1 Stratagene mouse testis (#937308) Mus

musculus cDNA clone 515187 5'). Based upon sequence similarity, dx290_1 proteins and each similar protein or peptide may share at least some activity.

Clone "ek390_4"

A polynucleotide of the present invention has been identified as clone "ek390_4". ek390_4 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. ek390_4 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "ek390_4 protein").

The nucleotide sequence of ek390_4 as presently determined is reported in SEQ ID NO:5. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the ek390_4 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:6. Amino acids 25 to 37 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 38, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone ek390_4 should be approximately 1000 bp.

The nucleotide sequence disclosed herein for ek390_4 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. ek390_4 demonstrated at least some similarity with sequences identified as AA075783 (zm89h02.r1 Stratagene ovarian cancer (#937219) Homo sapiens cDNA clone 545139 5'), AA427538 (zw32g04.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone 771030 5'), AA427539 (zw32g04.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone 771030 3'), AA453353 (zx47a06.r1 Soares testis NHT Homo sapiens cDNA clone 795346 5'), C20637 (HUMGS0004639, Human Gene Signature, 3'-directed cDNA sequence), R74326 (yl01c07.s1 Homo sapiens cDNA clone 156972 3'), R74420 (yl01c07.r1 Homo sapiens cDNA clone 156972 5'), T22914 (Human gene signature), U41197 (Human [TTTC]10 short tandem repeat polymorphism UM65, D17S1340), and X58237 (Human mRNA for anti-lectin antibody epitope (clone p36/8-6)). Based upon sequence similarity, ek390_4 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts a potential

transmembrane domain within the ek390_4 protein sequence centered around amino acid 160 of SEQ ID NO:6. The nucleotide sequence of ek390_4 indicates that it may contain GGGA repeat sequences.

Clone "er471_7"

A polynucleotide of the present invention has been identified as clone "er471_7". er471_7 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. er471_7 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "er471_7 protein").

The nucleotide sequence of er471_7 as presently determined is reported in SEQ ID NO:7. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the er471_7 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:8. Amino acids 74 to 86 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 87, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone er471_7 should be approximately 2250 bp.

The nucleotide sequence disclosed herein for er471_7 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. er471_7 demonstrated at least some similarity with sequences identified as AA039137 (mi98h06.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA clone 474683 5'), AA066962 (mm38g05.r1 Stratagene mouse melanoma (#937312) Mus musculus cDNA clone 523832 5'), AA189170 (zq47h05.s1 Stratagene hNT neuron (#937233) Homo sapiens cDNA clone 632889 3'), AA609188 (af12c10.s1 Soares testis NHT Homo sapiens cDNA clone 1031442 3'), and W07704 (zb02e02.r1 Soares fetal lung NbHL19W Homo sapiens cDNA clone 300890 5' similar to SW:YN66_YEAST P40164 HYPOTHETICAL 98.1 KD PROTEIN IN SPX19-GCR2 INTERGENIC REGION). The predicted amino acid sequence disclosed herein for er471_7 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted er471_7 protein demonstrated at least

some similarity to sequences identified as AF016448 (Cosmid F41E6 [Caenorhabditis elegans]) and L08407 (collagen type XVII [Mus musculus]). Based upon sequence similarity, er471_7 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts three potential transmembrane domains within the er471_7 protein sequence, centered around amino acids 40, 80, and 110 of SEQ ID NO:8, respectively.

Clone "fs40_3"

A polynucleotide of the present invention has been identified as clone "fs40_3". fs40_3 was isolated from a human adult testes cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. fs40_3 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "fs40_3 protein").

The nucleotide sequence of fs40_3 as presently determined is reported in SEQ ID NO:9. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the fs40_3 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:10.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone fs40_3 should be approximately 1000 bp.

The nucleotide sequence disclosed herein for fs40_3 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. fs40_3 demonstrated at least some similarity with sequences identified as AA411142 (zt37g01.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone 724560 5'), AA412527 (zu12a03.s1 Soares testis NHT Homo sapiens cDNA clone 731596 3'), AA565855 (nj32d09.s1 NCI_CGAP_AA1 Homo sapiens cDNA clone IMAGE:994193), H17042 (ym39f12.s1 Homo sapiens cDNA clone 50584 3'), and T33280 (EST57284 Homo sapiens cDNA 3' end similar to None). Based upon sequence similarity, fs40_3 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts a potential transmembrane domain within the fs40_3 protein sequence at the C-terminus of SEQ ID NO:10.

Clone "ga63_6"

A polynucleotide of the present invention has been identified as clone "ga63_6". ga63_6 was isolated from a human adult testes cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. ga63_6 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "ga63_6 protein").

The nucleotide sequence of ga63_6 as presently determined is reported in SEQ ID NO:11. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the ga63_6 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:12. Amino acids 11 to 23 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 24, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone ga63_6 should be approximately 2300 bp.

The nucleotide sequence disclosed herein for ga63_6 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. ga63_6 demonstrated at least some similarity with sequences identified as AA405433 (zu13h10.r1 Soares testis NHT Homo sapiens cDNA clone 731779 5'similar to TR G474970 G474970 SP32 PRECURSOR), AA406076 (zu67c02.s1 Soares testis NHT Homo sapiens cDNA clone 743042 3' similar to TR:G475021 G475021 SP32 PRECURSOR), AA424694 (zu13h10.s1 Soares testis NHT Homo sapiens cDNA clone 731779 3' similar to TR G475021 G475021 SP32 PRECURSOR; contains element TAR1 repetitive element), D16200 (Pig mRNA for sp32, partial sequence), D16203 (Guinea pig mRNA for sp32, complete cds), and D17573 (Mouse mRNA for proacrosin-binding protein (sp32), complete cds). The predicted amino acid sequence disclosed herein for ga63_6 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted ga63_6 protein demonstrated at least some similarity to sequences identified as D16200 (sp32 precursor [Sus scrofa]), and D17574 (alternative splicing product for proacrosin-binding protein (sp32) [Mus musculus]). The sp32 protein is found in the acrosomal vescicle of sperm, which is involved in egg-sperm fusion in fertilization. This protein is initially synthesized

as a 61-kDa precursor protein with a putative signal peptide at the amino terminus. The carboxyl-terminal half of the precursor molecule corresponds to the mature sp32 protein. Thus, sp32 is produced by post-translational modification of the precursor. The binding of sp32 to proacrosin may be involved in packaging the acrosin zymogen into the acrosomal matrix. (Baba *et al.*, 1994, *J. Biol. Chem.* **269 (13)**: 10133-10140, which is incorporated by reference herein). Based upon sequence similarity, ga63_6 proteins and each similar protein or peptide may share at least some activity.

Clone "gm335_4"

A polynucleotide of the present invention has been identified as clone "gm335_4". gm335_4 was isolated from a human adult uterus cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. gm335_4 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "gm335_4 protein").

The nucleotide sequence of gm335_4 as presently determined is reported in SEQ ID NO:13. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the gm335_4 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:14. Amino acids 8 to 20 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 21, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone gm335_4 should be approximately 800 bp.

The nucleotide sequence disclosed herein for gm335_4 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. gm335_4 demonstrated at least some similarity with sequences identified as AA055367 (zf20b05.r1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 377457 5'), AC002389 (Human DNA from chromosome 19 specific cosmid R28461, genomic sequence, complete sequence), W08522 (mb46h10.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 332515 5'), and X93916 (S.scrofa mRNA (clone VIB11; expressed sequence tag)). Based upon sequence similarity, gm335_4 proteins and each similar protein or peptide may share at least some activity.

Clone "hy370_9"

A polynucleotide of the present invention has been identified as clone "hy370_9". hy370_9 was isolated from a human adult trachea cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. hy370_9 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "hy370_9 protein").

The nucleotide sequence of hy370_9 as presently determined is reported in SEQ ID NO:15. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the hy370_9 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:16. Amino acids 8 to 20 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 21, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone hy370_9 should be approximately 1200 bp.

The nucleotide sequence disclosed herein for hy370_9 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. hy370_9 demonstrated at least some similarity with the sequence identified as AA763313 (vv89h07.r1 Stratagene mouse skin (#937313) Mus musculus cDNA clone 1229629 5'). Based upon sequence similarity, hy370_9 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts an additional potential transmembrane domain within the hy370_9 protein sequence centered around amino acid 140 of SEQ ID NO:16.

Clone "ie47_4"

A polynucleotide of the present invention has been identified as clone "ie47_4". ie47_4 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. ie47_4 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "ie47_4 protein").

The nucleotide sequence of ie47_4 as presently determined is reported in SEQ ID NO:17. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the ie47_4 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:18. Amino acids 17 to 29 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 30, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone ie47_4 should be approximately 2300 bp.

The nucleotide sequence disclosed herein for ie47_4 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. ie47_4 demonstrated at least some similarity with sequences identified as AA071953 (mf17h08.r1 Life Tech mouse brain Mus musculus cDNA clone 405375 5' similar to TR G304421 G304421 SILENCER ELEMENT), AA207250 (zq82d05.s1 Stratagene hNT neuron (#937233) Homo sapiens cDNA clone 648105 3' similar to TR G304421 G304421 SILENCER ELEMENT), L14938 (Chicken SCG10 protein mRNA, complete cds), L20260 (Mouse SCG10 gene sequence), R49053 (yg58c05.s1 Homo sapiens cDNA clone 37017 3'), S82024 (SCG10 neuron-specific growth-associated protein/stathmin homolog [human, embryo, mRNA]), T25428 (Human gene signature HUMGS07594, T25428 standard; cDNA to mRNA), W54204 (md04a12.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA clone 367390 5' similar to SW:SCGB_XENLA Q09002 SCG10 PROTEIN HOMOLOG A), X71433 (X. laevis SCG10 mRNA), and Z99916 (Human DNA sequence *** SEQUENCING IN PROGRESS *** from clone 221G9; HTGS phase 1). The predicted amino acid sequence disclosed herein for ie47_4 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted ie47_4 protein demonstrated at least some similarity to sequences identified as L14938 (SCG10 protein [Gallus gallus]) and S82024 (SCG10 neuron-specific growth-associated protein/stathmin homolog [human, embryo, Peptide] [Homo sapiens]). SCG10 protein is considered to be a membrane-bound protein present in neural growth cones and developing neurons (Maucuer *et al.*, 1993, *J. Biol. Chem.* **268**: 16420-16429; Stein *et al.*, 1988, *Neuron* **1**:463-476; which are incorporated by reference herein). Based upon sequence similarity, ie47_4 proteins and each similar protein or peptide may share at least some activity.

Clone "s195_10"

A polynucleotide of the present invention has been identified as clone "s195_10". s195_10 was isolated from a human adult neural tissue cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. s195_10 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "s195_10 protein").

The nucleotide sequence of s195_10 as presently determined is reported in SEQ ID NO:19. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the s195_10 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:20. Amino acids 35 to 47 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 48, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone s195_10 should be approximately 3500 bp.

The nucleotide sequence disclosed herein for s195_10 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. s195_10 demonstrated at least some similarity with sequences identified as AA113800 (zn65b05.s1 Stratagene HeLa cell s3 937216 Homo sapiens cDNA clone 563025 3' similar to TR:G600018 G600018 SSM4P), AA114062 (zn65b05.r1 Stratagene HeLa cell s3 937216 Homo sapiens cDNA clone 563025 5'), AA280316 (zt10f06.s1 Soares NbHTGBC Homo sapiens cDNA clone 712739 3'), AF009301 (Homo sapiens TEB4 protein mRNA, complete cds), N70344 (za60f10.s1 Homo sapiens cDNA clone 296971 3'), R60474 (yh13g07.r1 Homo sapiens cDNA clone 43058 5'), and T26266 (standard; cDNA to mRNA; 148 BP, Human gene signature HUMGS08505). The predicted amino acid sequence disclosed herein for s195_10 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted s195_10 protein demonstrated at least some similarity to sequences identified as AF009301 (TEB4 protein [Homo sapiens]), X76715 (SSM4 gene product [Saccharomyces cerevisiae]), Z46861 (Ssm4p [Saccharomyces cerevisiae]), and Z47047 (Ssm4p [Saccharomyces cerevisiae]). Based upon sequence similarity, s195_10 proteins and each similar protein or peptide may share at least some activity. The TopPredII

computer program predicts eleven additional potential transmembrane domains within the s195_10 protein sequence, centered around amino acids 130, 170, 210, 260, 320, 470, 520, 560, 600, 650, and 690 of SEQ ID NO:20, respectively. The nucleotide sequence of s195_10 indicates that it may contain a simple GAA repeat region.

Deposit of Clones

Clones do15_4, dx290_1, ek390_4, er471_7, fs40_3, ga63_6, gm335_4, hy370_9, ie47_4, and s195_10 were deposited on June 19, 1997 with the American Type Culture Collection (10801 University Boulevard, Manassas, Virginia 20110-2209 U.S.A.) as an original deposit under the Budapest Treaty and were given the accession number ATCC 98468, from which each clone comprising a particular polynucleotide is obtainable. All restrictions on the availability to the public of the deposited material will be irrevocably removed upon the granting of the patent, except for the requirements specified in 37 C.F.R. § 1.808(b), and the term of the deposit will comply with 37 C.F.R. § 1.806.

Each clone has been transfected into separate bacterial cells (*E. coli*) in this composite deposit. Each clone can be removed from the vector in which it was deposited by performing an EcoRI/NotI digestion (5' site, EcoRI; 3' site, NotI) to produce the appropriate fragment for such clone. Each clone was deposited in either the pED6 or pNOTs vector depicted in Figures 1A and 1B, respectively. The pED6dpc2 vector ("pED6") was derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning (Kaufman *et al.*, 1991, *Nucleic Acids Res.* **19**: 4485-4490); the pNOTs vector was derived from pMT2 (Kaufman *et al.*, 1989, *Mol. Cell. Biol.* 9: 946-958) by deletion of the DHFR sequences, insertion of a new polylinker, and insertion of the M13 origin of replication in the ClaI site. In some instances, the deposited clone can become "flipped" (i.e., in the reverse orientation) in the deposited isolate. In such instances, the cDNA insert can still be isolated by digestion with EcoRI and NotI. However, NotI will then produce the 5' site and EcoRI will produce the 3' site for placement of the cDNA in proper orientation for expression in a suitable vector. The cDNA may also be expressed from the vectors in which they were deposited.

Bacterial cells containing a particular clone can be obtained from the composite deposit as follows:

An oligonucleotide probe or probes should be designed to the sequence that is known for that particular clone. This sequence can be derived from the sequences provided herein, or from a combination of those sequences. The sequence of an

oligonucleotide probe that was used to isolate or to sequence each full-length clone is identified below, and should be most reliable in isolating the clone of interest.

Clone	Probe Sequence
do15_4	SEQ ID NO:21
dx290_1	SEQ ID NO:22
ek390_4	SEQ ID NO:23
er471_7	SEQ ID NO:24
fs40_3	SEQ ID NO:25
ga63_6	SEQ ID NO:26
gm335_4	SEQ ID NO:27
hy370_9	SEQ ID NO:28
ie47_4	SEQ ID NO:29
s195_10	SEQ ID NO:30

In the sequences listed above which include an N at position 2, that position is occupied in preferred probes/primers by a biotinylated phosphoaramidite residue rather than a nucleotide (such as , for example, that produced by use of biotin phosphoramidite (1-dimethoxytrityloxy-2-(N-biotinyl-4-aminobutyl)-propyl-3-O-(2-cyanoethyl)-(N,N-diisopropyl)-phosphoramadite) (Glen Research, cat. no. 10-1953)).

The design of the oligonucleotide probe should preferably follow these parameters:

(a) It should be designed to an area of the sequence which has the fewest ambiguous bases ("N's"), if any;

(b) It should be designed to have a T_m of approx. 80 ° C (assuming 2° for each A or T and 4 degrees for each G or C).

The oligonucleotide should preferably be labeled with g-^{32}P ATP (specific activity 6000 Ci/mmole) and T4 polynucleotide kinase using commonly employed techniques for labeling oligonucleotides. Other labeling techniques can also be used. Unincorporated label should preferably be removed by gel filtration chromatography or other established methods. The amount of radioactivity incorporated into the probe should be quantitated by measurement in a scintillation counter. Preferably, specific activity of the resulting probe should be approximately 4e+6 dpm/pmole.

The bacterial culture containing the pool of full-length clones should preferably be thawed and 100 µl of the stock used to inoculate a sterile culture flask containing 25 ml of sterile L-broth containing ampicillin at 100 µg/ml. The culture should preferably be grown to saturation at 37°C, and the saturated culture should preferably be diluted in fresh L-broth. Aliquots of these dilutions should preferably be plated to determine the dilution and volume which will yield approximately 5000 distinct and well-separated colonies on solid bacteriological media containing L-broth containing ampicillin at 100 µg/ml and agar at 1.5% in a 150 mm petri dish when grown overnight at 37°C. Other known methods of obtaining distinct, well-separated colonies can also be employed.

Standard colony hybridization procedures should then be used to transfer the colonies to nitrocellulose filters and lyse, denature and bake them.

The filter is then preferably incubated at 65°C for 1 hour with gentle agitation in 6X SSC (20X stock is 175.3 g NaCl/liter, 88.2 g Na citrate/liter, adjusted to pH 7.0 with NaOH) containing 0.5% SDS, 100 µg/ml of yeast RNA, and 10 mM EDTA (approximately 10 mL per 150 mm filter). Preferably, the probe is then added to the hybridization mix at a concentration greater than or equal to 1e+6 dpm/mL. The filter is then preferably incubated at 65°C with gentle agitation overnight. The filter is then preferably washed in 500 mL of 2X SSC/0.5% SDS at room temperature without agitation, preferably followed by 500 mL of 2X SSC/0.1% SDS at room temperature with gentle shaking for 15 minutes. A third wash with 0.1X SSC/0.5% SDS at 65°C for 30 minutes to 1 hour is optional. The filter is then preferably dried and subjected to autoradiography for sufficient time to visualize the positives on the X-ray film. Other known hybridization methods can also be employed.

The positive colonies are picked, grown in culture, and plasmid DNA isolated using standard procedures. The clones can then be verified by restriction analysis, hybridization analysis, or DNA sequencing.

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H.U. Saragovi, *et al.*, Bio/Technology 10, 773-778 (1992) and in R.S. McDowell, *et al.*, J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites. For example, fragments of the protein may be fused through "linker" sequences to

the Fc portion of an immunoglobulin. For a bivalent form of the protein, such a fusion could be to the Fc portion of an IgG molecule. Other immunoglobulin isotypes may also be used to generate such fusions. For example, a protein - IgM fusion would generate a decavalent form of the protein of the invention.

The present invention also provides both full-length and mature forms of the disclosed proteins. The full-length form of the such proteins is identified in the sequence listing by translation of the nucleotide sequence of each disclosed clone. The mature form(s) of such protein may be obtained by expression of the disclosed full-length polynucleotide (preferably those deposited with ATCC) in a suitable mammalian cell or other host cell. The sequence(s) of the mature form(s) of the protein may also be determinable from the amino acid sequence of the full-length form.

The present invention also provides genes corresponding to the polynucleotide sequences disclosed herein. "Corresponding genes" are the regions of the genome that are transcribed to produce the mRNAs from which cDNA polynucleotide sequences are derived and may include contiguous regions of the genome necessary for the regulated expression of such genes. Corresponding genes may therefore include but are not limited to coding sequences, 5' and 3' untranslated regions, alternatively spliced exons, introns, promoters, enhancers, and silencer or suppressor elements. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. An "isolated gene" is a gene that has been separated from the adjacent coding sequences, if any, present in the genome of the organism from which the gene was isolated.

Organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein are provided. The desired change in gene expression can be achieved through the use of antisense polynucleotides or ribozymes that bind and/or cleave the mRNA transcribed from the gene (Albert and Morris, 1994, *Trends Pharmacol. Sci.* **15(7)**: 250-254; Lavarosky *et al.*, 1997, *Biochem. Mol. Med.* **62(1)**: 11-22; and Hampel, 1998, *Prog. Nucleic Acid Res. Mol. Biol.* **58**: 1-39; all of which are incorporated by reference herein). Transgenic animals that have multiple copies of the gene(s) corresponding to the polynucleotide sequences disclosed herein, preferably produced by transformation of cells with genetic constructs that are stably maintained within the transformed cells and their progeny, are provided.

Transgenic animals that have modified genetic control regions that increase or reduce gene expression levels, or that change temporal or spatial patterns of gene expression, are also provided (see European Patent No. 0 649 464 B1, incorporated by reference herein). In addition, organisms are provided in which the gene(s) corresponding to the polynucleotide sequences disclosed herein have been partially or completely inactivated, through insertion of extraneous sequences into the corresponding gene(s) or through deletion of all or part of the corresponding gene(s). Partial or complete gene inactivation can be accomplished through insertion, preferably followed by imprecise excision, of transposable elements (Plasterk, 1992, *Bioessays* **14(9)**: 629-633; Zwaal *et al.*, 1993, *Proc. Natl. Acad. Sci. USA* **90(16)**: 7431-7435; Clark *et al.*, 1994, *Proc. Natl. Acad. Sci. USA* **91(2)**: 719-722; all of which are incorporated by reference herein), or through homologous recombination, preferably detected by positive/negative genetic selection strategies (Mansour *et al.*, 1988, *Nature* **336**: 348-352; U.S. Patent Nos. 5,464,764; 5,487,992; 5,627,059; 5,631,153; 5,614, 396; 5,616,491; and 5,679,523; all of which are incorporated by reference herein). These organisms with altered gene expression are preferably eukaryotes and more preferably are mammals. Such organisms are useful for the development of non-human models for the study of disorders involving the corresponding gene(s), and for the development of assay systems for the identification of molecules that interact with the protein product(s) of the corresponding gene(s).

Where the protein of the present invention is membrane-bound (e.g., is a receptor), the present invention also provides for soluble forms of such protein. In such forms part or all of the intracellular and transmembrane domains of the protein are deleted such that the protein is fully secreted from the cell in which it is expressed. The intracellular and transmembrane domains of proteins of the invention can be identified in accordance with known techniques for determination of such domains from sequence information.

Proteins and protein fragments of the present invention include proteins with amino acid sequence lengths that are at least 25%(more preferably at least 50%, and most preferably at least 75%) of the length of a disclosed protein and have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with that disclosed protein, where sequence identity is determined by comparing the amino acid sequences of the proteins when aligned so as to maximize overlap and identity while minimizing sequence gaps. Also included in the present invention are proteins and protein fragments that contain a segment preferably comprising 8 or more (more preferably 20 or more, most preferably 30 or more) contiguous amino acids that

shares at least 75% sequence identity (more preferably, at least 85% identity; most preferably at least 95% identity) with any such segment of any of the disclosed proteins.

Species homologues of the disclosed polynucleotides and proteins are also provided by the present invention. As used herein, a "species homologue" is a protein or polynucleotide with a different species of origin from that of a given protein or polynucleotide, but with significant sequence similarity to the given protein or polynucleotide. Preferably, polynucleotide species homologues have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% identity) with the given polynucleotide, and protein species homologues have at least 30% sequence identity (more preferably, at least 45% identity; most preferably at least 60% identity) with the given protein, where sequence identity is determined by comparing the nucleotide sequences of the polynucleotides or the amino acid sequences of the proteins when aligned so as to maximize overlap and identity while minimizing sequence gaps. Species homologues may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species. Preferably, species homologues are those isolated from mammalian species. Most preferably, species homologues are those isolated from certain mammalian species such as, for example, *Pan troglodytes, Gorilla gorilla, Pongo pygmaeus, Hylobates concolor, Macaca mulatta, Papio papio, Papio hamadryas, Cercopithecus aethiops, Cebus capucinus, Aotus trivirgatus, Sanguinus oedipus, Microcebus murinus, Mus musculus, Rattus norvegicus, Cricetulus griseus, Felis catus, Mustela vison, Canis familiaris, Oryctolagus cuniculus, Bos taurus, Ovis aries, Sus scrofa,* and *Equus caballus,* for which genetic maps have been created allowing the identification of syntenic relationships between the genomic organization of genes in one species and the genomic organization of the related genes in another species (O'Brien and Seuánez, 1988, *Ann. Rev. Genet.* **22**: 323-351; O'Brien *et al.*, 1993, *Nature Genetics* **3**:103-112; Johansson *et al.*, 1995, *Genomics* **25**: 682-690; Lyons *et al.*, 1997, *Nature Genetics* **15**: 47-56; O'Brien *et al.*, 1997, *Trends in Genetics* **13(10)**: 393-399; Carver and Stubbs, 1997, *Genome Research* **7**:1123-1137; all of which are incorporated by reference herein).

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotides which also encode proteins which are identical or have significantly similar sequences to those encoded by the disclosed polynucleotides. Preferably, allelic variants have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90%

identity) with the given polynucleotide, where sequence identity is determined by comparing the nucleotide sequences of the polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps. Allelic variants may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from individuals of the appropriate species.

The invention also includes polynucleotides with sequences complementary to those of the polynucleotides disclosed herein.

The present invention also includes polynucleotides that hybridize under reduced stringency conditions, more preferably stringent conditions, and most preferably highly stringent conditions, to polynucleotides described herein. Examples of stringency conditions are shown in the table below: highly stringent conditions are those that are at least as stringent as, for example, conditions A-F; stringent conditions are at least as stringent as, for example, conditions G-L; and reduced stringency conditions are at least as stringent as, for example, conditions M-R.

Stringency Condition	Polynucleotide Hybrid	Hybrid Length (bp)‡	Hybridization Temperature and Buffer†	Wash Temperature and Buffer†
A	DNA:DNA	≥ 50	65°C; 1xSSC -or- 42°C; 1xSSC, 50% formamide	65°C; 0.3xSSC
B	DNA:DNA	<50	T_B*; 1xSSC	T_B*; 1xSSC
C	DNA:RNA	≥ 50	67°C; 1xSSC -or- 45°C; 1xSSC, 50% formamide	67°C; 0.3xSSC
D	DNA:RNA	<50	T_D*; 1xSSC	T_D*; 1xSSC
E	RNA:RNA	≥ 50	70°C; 1xSSC -or- 50°C; 1xSSC, 50% formamide	70°C; 0.3xSSC
F	RNA:RNA	<50	T_F*; 1xSSC	T_F*; 1xSSC
G	DNA:DNA	≥ 50	65°C; 4xSSC -or- 42°C; 4xSSC, 50% formamide	65°C; 1xSSC
H	DNA:DNA	<50	T_H*; 4xSSC	T_H*; 4xSSC
I	DNA:RNA	≥ 50	67°C; 4xSSC -or- 45°C; 4xSSC, 50% formamide	67°C; 1xSSC
J	DNA:RNA	<50	T_J*; 4xSSC	T_J*; 4xSSC
K	RNA:RNA	≥ 50	70°C; 4xSSC -or- 50°C; 4xSSC, 50% formamide	67°C; 1xSSC
L	RNA:RNA	<50	T_L*; 2xSSC	T_L*; 2xSSC
M	DNA:DNA	≥ 50	50°C; 4xSSC -or- 40°C; 6xSSC, 50% formamide	50°C; 2xSSC
N	DNA:DNA	<50	T_N*; 6xSSC	T_N*; 6xSSC
O	DNA:RNA	≥ 50	55°C; 4xSSC -or- 42°C; 6xSSC, 50% formamide	55°C; 2xSSC
P	DNA:RNA	<50	T_P*; 6xSSC	T_P*; 6xSSC
Q	RNA:RNA	≥ 50	60°C; 4xSSC -or- 45°C; 6xSSC, 50% formamide	60°C; 2xSSC
R	RNA:RNA	<50	T_R*; 4xSSC	T_R*; 4xSSC

‡: The hybrid length is that anticipated for the hybridized region(s) of the hybridizing polynucleotides. When hybridizing a polynucleotide to a target polynucleotide of unknown sequence, the hybrid length is assumed to be that of the hybridizing polynucleotide. When polynucleotides of known sequence are hybridized, the hybrid length can be determined by aligning the sequences of the polynucleotides and identifying the region or regions of optimal sequence complementarity.

†: SSPE (1xSSPE is 0.15M NaCl, 10mM NaH_2PO_4, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes after hybridization is complete.

*T_B - T_R: The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T_m) of the hybrid, where T_m is determined according to the following equations. For hybrids less than 18 base pairs in length, T_m(°C) = 2(# of A + T bases) + 4(# of G + C bases). For hybrids between 18 and 49 base pairs in length, T_m(°C) = 81.5 + 16.6($\log_{10}$[Na^+]) + 0.41(%G+C) - (600/N), where N is the number of bases in the hybrid, and [Na^+] is the concentration of sodium ions in the hybridization buffer ([Na^+] for 1xSSC = 0.165 M).

Additional examples of stringency conditions for polynucleotide hybridization are provided in Sambrook, J., E.F. Fritsch, and T. Maniatis, 1989, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, chapters 9 and 11, and *Current Protocols in Molecular Biology*, 1995, F.M. Ausubel et al., eds., John Wiley & Sons, Inc., sections 2.10 and 6.3-6.4, incorporated herein by reference.

Preferably, each such hybridizing polynucleotide has a length that is at least 25%(more preferably at least 50%, and most preferably at least 75%) of the length of the polynucleotide of the present invention to which it hybridizes, and has at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with the polynucleotide of the present invention to which it hybridizes, where sequence identity is determined by comparing the sequences of the hybridizing polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps.

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman *et al.*, Nucleic Acids Res. 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, Methods in Enzymology 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

A number of types of cells may act as suitable host cells for expression of the protein. Mammalian host cells include, for example, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from in vitro culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or in prokaryotes such as bacteria. Potentially suitable yeast strains include *Saccharomyces cerevisiae, Schizosaccharomyces pombe, Kluyveromyces* strains, *Candida*, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli, Bacillus subtilis, Salmonella typhimurium*, or any bacterial

strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *e.g.*, Invitrogen, San Diego, California, U.S.A. (the MaxBac® kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (i.e., from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl® or Cibacrom blue 3GA Sepharose®; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX). Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, MA), Pharmacia (Piscataway, NJ) and InVitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("Flag") is commercially available from Kodak (New Haven, CT).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant

methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The protein may also be produced by known conventional chemical synthesis. Methods for constructing the proteins of the present invention by synthetic means are known to those skilled in the art. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications in the peptide or DNA sequences can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, e.g., U.S. Patent No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and may thus be useful for screening or other immunological methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are believed to be encompassed by the present invention.

USES AND BIOLOGICAL ACTIVITY

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified below. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or by administration or use of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA).

Research Uses and Utilities

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on Southern gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, those described in Gyuris *et al.*, 1993, *Cell* 75: 791-803 and in Rossi *et al.*, 1997, *Proc. Natl. Acad. Sci. USA* 94: 8405-8410, all of which are incorporated by reference herein) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The proteins provided by the present invention can similarly be used in assay to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine

levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Where the protein binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the protein can be used to identify the other protein with which binding occurs or to identify inhibitors of the binding interaction. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E.F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S.L. and A.R. Kimmel eds., 1987.

Nutritional Uses

Polynucleotides and proteins of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the protein or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the protein or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

Cytokine and Cell Proliferation/Differentiation Activity

A protein of the present invention may exhibit cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of a protein of the present invention is

evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+ (preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e and CMK.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., J. Immunol. 149:3778-3783, 1992; Bowman et al., J. Immunol. 152: 1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A.M. and Shevach, E.M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human Interferon γ, Schreiber, R.D. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L.S. and Lipsky, P.E. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse and human interleukin 6 - Nordan, R. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Acad. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11 - Bennett, F., Giannotti, J., Clark, S.C. and Turner, K. J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9 - Ciarletta, A., Giannotti, J., Clark, S.C. and Turner, K.J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

Immune Stimulating or Suppressing Activity

A protein of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases causes by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpesviruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, a protein of the present invention may also be useful where a boost to the immune system generally may be desirable, *i.e.*, in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein of the present invention may also to be useful in the treatment of allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for

example, organ transplantation), may also be treatable using a protein of the present invention.

Using the proteins of the invention it may also be possible to immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as , for example, B7)), *e.g.*, preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a molecule which inhibits or blocks interaction of a B7 lymphocyte antigen with its natural ligand(s) on immune cells (such as a soluble, monomeric form of a peptide having B7-2 activity alone or in conjunction with a monomeric form of a peptide having an activity of another B lymphocyte antigen (*e.g.*, B7-1, B7-3) or blocking antibody), prior to transplantation can lead to the binding of the molecule to the natural ligand(s) on the immune cells without transmitting the corresponding costimulatory signal. Blocking B lymphocyte antigen function in this matter prevents cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, the lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or

tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular blocking reagents in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins *in vivo* as described in Lenschow *et al.*, Science 257:789-792 (1992) and Turka *et al.*, Proc. Natl. Acad. Sci USA, *89*:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of blocking B lymphocyte antigen function *in vivo* on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block costimulation of T cells by disrupting receptor:ligand interactions of B lymphocyte antigens can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/*lpr/lpr* mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (preferably a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response through stimulating B lymphocyte antigen function may be useful in cases of

viral infection. In addition, systemic viral diseases such as influenza, the common cold, and encephalitis might be alleviated by the administration of stimulatory forms of B lymphocyte antigens systemically.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells *in vitro* with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the *in vitro* activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells *in vivo*.

In another application, up regulation or enhancement of antigen function (preferably B lymphocyte antigen function) may be useful in the induction of tumor immunity. Tumor cells (*e.g.*, sarcoma, melanoma, lymphoma, leukemia, neuroblastoma, carcinoma) transfected with a nucleic acid encoding at least one peptide of the present invention can be administered to a subject to overcome tumor-specific tolerance in the subject. If desired, the tumor cell can be transfected to express a combination of peptides. For example, tumor cells obtained from a patient can be transfected *ex vivo* with an expression vector directing the expression of a peptide having B7-2-like activity alone, or in conjunction with a peptide having B7-1-like activity and/or B7-3-like activity. The transfected tumor cells are returned to the patient to result in expression of the peptides on the surface of the transfected cell. Alternatively, gene therapy techniques can be used to target a tumor cell for transfection *in vivo*.

The presence of the peptide of the present invention having the activity of a B lymphocyte antigen(s) on the surface of the tumor cell provides the necessary costimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient amounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (*e.g.*, a cytoplasmic-domain truncated portion) of an MHC class I α chain protein and β_2 microglobulin protein or an MHC class II α chain protein and an MHC class II β chain protein to thereby express MHC class I or MHC class II proteins on the cell surface.

Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (*e.g.*, B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Bowmanet al., J. Virology 61:1992-1998; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: *In vitro* antibody production, Mond, J.J. and Brunswick, M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-

Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

Hematopoiesis Regulating Activity

A protein of the present invention may be useful in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell deficiencies. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity)

useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either *in-vivo* or *ex-vivo* (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M.G. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, NY. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I.K. and Briddell, R.A. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, NY. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R.E. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 1-21, Wiley-Liss, Inc.., New York, NY. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 163-179,

Wiley-Liss, Inc., New York, NY. 1994; Long term culture initiating cell assay, Sutherland, H.J. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, NY. 1994.

Tissue Growth Activity

A protein of the present invention also may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as for wound healing and tissue repair and replacement, and in the treatment of burns, incisions and ulcers.

A protein of the present invention, which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Such a preparation employing a protein of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. *De novo* bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A protein of this invention may also be used in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce differentiation of progenitors of bone-forming cells. A protein of the invention may also be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes.

Another category of tissue regeneration activity that may be attributable to the protein of the present invention is tendon/ligament formation. A protein of the present invention, which induces tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue

formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide an environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The protein of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a protein may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

Proteins of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

It is expected that a protein of the present invention may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring to allow normal tissue to regenerate. A protein of the invention may also exhibit angiogenic activity.

A protein of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A protein of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, HI and Rovee, DT, eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

Activin/Inhibin Activity

A protein of the present invention may also exhibit activin- or inhibin-related activities. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a protein of the present invention, alone or in heterodimers with a member of the inhibin α family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the protein of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin-β group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, United States Patent 4,798,885. A protein of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as cows, sheep and pigs.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

Chemotactic/Chemokinetic Activity

A protein of the present invention may have chemotactic or chemokinetic activity (e.g., act as a chemokine) for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. Chemotactic and chemokinetic proteins can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic proteins provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W.Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25: 1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153: 1762-1768, 1994.

Hemostatic and Thrombolytic Activity

A protein of the invention may also exhibit hemostatic or thrombolytic activity. As a result, such a protein is expected to be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A protein of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

Receptor/Ligand Activity

A protein of the present invention may also demonstrate activity as receptors, receptor ligands or inhibitors or agonists of receptor/ligand interactions. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses). Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in:Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W.Strober, Pub. Greene Publishing Associates and

Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

Anti-Inflammatory Activity

Proteins of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Proteins exhibiting such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation inflammation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Proteins of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material.

Cadherin/Tumor Invasion Suppressor Activity

Cadherins are calcium-dependent adhesion molecules that appear to play major roles during development, particularly in defining specific cell types. Loss or alteration of normal cadherin expression can lead to changes in cell adhesion properties linked to tumor growth and metastasis. Cadherin malfunction is also implicated in other human diseases, such as pemphigus vulgaris and pemphigus foliaceus (auto-immune blistering skin diseases), Crohn's disease, and some developmental abnormalities.

The cadherin superfamily includes well over forty members, each with a distinct pattern of expression. All members of the superfamily have in common conserved extracellular repeats (cadherin domains), but structural differences are found in other parts of the molecule. The cadherin domains bind calcium to form their tertiary structure and thus calcium is required to mediate their adhesion. Only a few amino acids in the

first cadherin domain provide the basis for homophilic adhesion; modification of this recognition site can change the specificity of a cadherin so that instead of recognizing only itself, the mutant molecule can now also bind to a different cadherin. In addition, some cadherins engage in heterophilic adhesion with other cadherins.

E-cadherin, one member of the cadherin superfamily, is expressed in epithelial cell types. Pathologically, if E-cadherin expression is lost in a tumor, the malignant cells become invasive and the cancer metastasizes. Transfection of cancer cell lines with polynucleotides expressing E-cadherin has reversed cancer-associated changes by returning altered cell shapes to normal, restoring cells' adhesiveness to each other and to their substrate, decreasing the cell growth rate, and drastically reducing anchorage-independent cell growth. Thus, reintroducing E-cadherin expression reverts carcinomas to a less advanced stage. It is likely that other cadherins have the same invasion suppressor role in carcinomas derived from other tissue types. Therefore, proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be used to treat cancer. Introducing such proteins or polynucleotides into cancer cells can reduce or eliminate the cancerous changes observed in these cells by providing normal cadherin expression.

Cancer cells have also been shown to express cadherins of a different tissue type than their origin, thus allowing these cells to invade and metastasize in a different tissue in the body. Proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be substituted in these cells for the inappropriately expressed cadherins, restoring normal cell adhesive properties and reducing or eliminating the tendency of the cells to metastasize.

Additionally, proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can used to generate antibodies recognizing and binding to cadherins. Such antibodies can be used to block the adhesion of inappropriately expressed tumor-cell cadherins, preventing the cells from forming a tumor elsewhere. Such an anti-cadherin antibody can also be used as a marker for the grade, pathological type, and prognosis of a cancer, i.e. the more progressed the cancer, the less cadherin expression there will be, and this decrease in cadherin expression can be detected by the use of a cadherin-binding antibody.

Fragments of proteins of the present invention with cadherin activity, preferably a polypeptide comprising a decapeptide of the cadherin recognition site, and polynucleotides of the present invention encoding such protein fragments, can also be used

to block cadherin function by binding to cadherins and preventing them from binding in ways that produce undesirable effects. Additionally, fragments of proteins of the present invention with cadherin activity, preferably truncated soluble cadherin fragments which have been found to be stable in the circulation of cancer patients, and polynucleotides encoding such protein fragments, can be used to disturb proper cell-cell adhesion.

Assays for cadherin adhesive and invasive suppressor activity include, without limitation, those described in: Hortsch et al. J Biol Chem 270 (32): 18809-18817, 1995; Miyaki et al. Oncogene 11: 2547-2552, 1995; Ozawa et al. Cell 63: 1033-1038, 1990.

Tumor Inhibition Activity

In addition to the activities described above for immunological treatment or prevention of tumors, a protein of the invention may exhibit other anti-tumor activities. A protein may inhibit tumor growth directly or indirectly (such as, for example, via ADCC). A protein may exhibit its tumor inhibitory activity by acting on tumor tissue or tumor precursor tissue, by inhibiting formation of tissues necessary to support tumor growth (such as, for example, by inhibiting angiogenesis), by causing production of other factors, agents or cell types which inhibit tumor growth, or by suppressing, eliminating or inhibiting factors, agents or cell types which promote tumor growth.

Other Activities

A protein of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or caricadic cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic

lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

ADMINISTRATION AND DOSING

A protein of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources) may be used in a pharmaceutical composition when combined with a pharmaceutically acceptable carrier. Such a composition may also contain (in addition to protein and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or compliment its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein of the invention, or to minimize side effects. Conversely, protein of the present invention may be included in formulations of the particular cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent.

A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunolgobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithin, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent No. 4,235,871; U.S. Patent No. 4,501,728; U.S. Patent No. 4,837,028; and U.S. Patent No. 4,737,323, all of which are incorporated herein by reference.

As used herein, the term "therapeutically effective amount" means the total amount of each active component of the pharmaceutical composition or method that is sufficient to show a meaningful patient benefit, i.e., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, the term refers to that ingredient alone. When applied to a combination, the term refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein of the present invention is administered to a mammal having a condition to be treated. Protein of the present invention may be

administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co-administered with one or more cytokines, lymphokines or other hematopoietic factors, protein of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

Administration of protein of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

When a therapeutically effective amount of protein of the present invention is administered orally, protein of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein of the present invention, and preferably from about 25 to 90% protein of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein of the present invention, and preferably from about 1 to 50% protein of the present invention.

When a therapeutically effective amount of protein of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred

pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art.

The amount of protein of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein of the present invention and observe the patient's response. Larger doses of protein of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 µg to about 100 mg (preferably about 0.1ng to about 10 mg, more preferably about 0.1 µg to about 1 mg) of protein of the present invention per kg body weight.

The duration of intravenous therapy using the pharmaceutical composition of the present invention will vary, depending on the severity of the disease being treated and the condition and potential idiosyncratic response of each individual patient. It is contemplated that the duration of each application of the protein of the present invention will be in the range of 12 to 24 hours of continuous intravenous administration. Ultimately the attending physician will decide on the appropriate duration of intravenous therapy using the pharmaceutical composition of the present invention.

Protein of the invention may also be used to immunize animals to obtain polyclonal and monoclonal antibodies which specifically react with the protein. Such antibodies may be obtained using either the entire protein or fragments thereof as an immunogen. The peptide immunogens additionally may contain a cysteine residue at the carboxyl terminus, and are conjugated to a hapten such as keyhole limpet hemocyanin (KLH). Methods for synthesizing such peptides are known in the art, for example, as in R.P. Merrifield, J. Amer.Chem.Soc. 85, 2149-2154 (1963); J.L. Krstenansky, *et al.*, FEBS Lett. 211, 10 (1987). Monoclonal antibodies binding to the protein of the invention may be useful diagnostic agents for the immunodetection of the protein. Neutralizing monoclonal

antibodies binding to the protein may also be useful therapeutics for both conditions associated with the protein and also in the treatment of some forms of cancer where abnormal expression of the protein is involved. In the case of cancerous cells or leukemic cells, neutralizing monoclonal antibodies against the protein may be useful in detecting and preventing the metastatic spread of the cancerous cells, which may be mediated by the protein.

For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalciumphosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalciumphosphate. The bioceramics may be altered in composition, such as in calcium-

aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability.

Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt%, preferably 1-10 wt% based on total formulation weight, which represents the amount necessary to prevent desorbtion of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells.

In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired patients for such treatment with proteins of the present invention.

The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, e.g., amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (e.g., bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in

the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either *in vivo* or *ex vivo* into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA).

Cells may also be cultured *ex vivo* in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes.

Patent and literature references cited herein are incorporated by reference as if fully set forth.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Jacobs, Kenneth
McCoy, John M.
LaVallie, Edward R.
Racie, Lisa A.
Treacy, Maurice
Spaulding, Vikki
Agostino, Michael J.
Howes, Steven H.
Fechtel, Kim

(ii) TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(iii) NUMBER OF SEQUENCES: 30

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Genetics Institute, Inc.
(B) STREET: 87 CambridgePark Drive
(C) CITY: Cambridge
(D) STATE: MA
(E) COUNTRY: U.S.A.
(F) ZIP: 02140

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Sprunger, Suzanne A.
(B) REGISTRATION NUMBER: 41,323

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (617) 498-8284
(B) TELEFAX: (617) 876-5851

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1748 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GTTTAGTGAT ACGACACAAG ATCGGGAGAT TTTTGATCAC CATACTGAAG AGGATATAGA 60

TAAAAGTGCT AACAGTGTAT TGATAAAAAA CCTGAGCAGG ACCCCATCTA GTTGCAGCAG 120

CTCTCTGGAT TCAATCAAGG CTGATGGGAC CTCTCTGGAC TTCAGCACTT ACCGCAGTAG 180

TCAAATGGAA TCACAGTTTC TCAGAGATAC TATTTGTGAA GAGAGCTTGA GGGAGAAACT 240

CCAAGATGGG AGAATAACAA TAAGGGAGTT CTTTATACTT CTCCAGGTCC ACATCTTGAT 300

ACAGAAACCC CGACAGAGCA ATCTCCCAGG CAATTTTACT GTAAACACAC CACCTACTCC 360

AGAAGACCTG ATGTTAAGTC AATATGTTTA CCGACCCAAG ATACAGATTT ATAGAGAAGA 420

TTGTGAGGCT CGTCGCCAAA AGATTGAAGA ATTAAAGCTT TCTGCATCGA ACCAAGATAA 480

GCTGTTGGTT GATATAAATA AGAACCTGTG GGAAAAAATG AGACACTGCT CTGACAAAGA 540

GCTGAAGGCC TTTGGAATTT ATCTTAACAA AATAAAGTCA TGTTTTACCA AGATGACTAA 600

AGTCTTCACT CACCAAGGAA AAGTGGCTCT GTATGGCAAG CTGGTGCAGT CAGCTCAGAA 660

TGAGAGGGAG AAACTTCAAA TAAAGATAGA TGAGATGGAT AAAATACTTA AGAAGATCGA 720

TAACTGCCTC ACTGAGATGG AAACAGAAAC TAAGAATTTG GAGGATGAAG AGAAAAACAA 780

TCCTGTGGAA GAATGGGATT CTGAAATGAG AGCTGCAGAA AAAGAATTGG AACAGCTGAA 840

AACTGAAGAG GAGGAGCTTC AAAGAAATCT CTTAGAACTG GAGGTACCAA AAGAGCAGAC 900

CCTTGCTCAA ATAGACTTTA TGCAAAAACA AAGAAATAGA ACTGAAGAGC TACTGGATCA 960

GTTGAGCTTG TCTGAGTGGG ATGTCGTTGA GTGGAGTGAT GATCAAGCTG TATTCACCTT 1020

TGTTTATGAC ACGATACAAC TCACCATCAC CTTTGAAGAG TCAGTTGTTG GTTTCCCTTT 1080

CCTGGACAAG CGTTATAGGA AGATTGTTGA TGTCAATTTT CAATCTCTGT TAGATGAGGA 1140

TCAAGCTCCT CCTTCCTCCC TTTTAGTTCA TAAGCTTATT TTCCAGTACG TTGAAGAAAA 1200

GGAATCCTGG AAGAAGACAT GTACAACCCA GCATCAGTTA CCCAAGATGC TTGAAGAATT 1260

CTCACTGGTA GTGCACCATT GCAGACTCCT TGGAGAGGAG ATTGAGTATT TAAAGAGATG 1320

GGGACCAAAT TATAACCTAA TGAACATAGA TATTAATAAT AATGAATTGA GACTTTTATT 1380

CTCTAGCTCC GCAGCATTTG CAAAGTTTGA AATAACTTTG TTTCTCTCAG CCTATTATCC 1440

ATCTGTACCA TTACCTTCCA CCATTCAGAA TCACGTTGGG AACACTAGCC AAGATGATAT 1500

TGCTACCATT CTATCTAAAG TGCCACTGGA GAACAACTAC CTGAAGAATG TAGTCAAGCA 1560

AATTTACCAA GATCTGTTTC AGGACTGCCA TTTCTACCAC TAGACCCTTG GACCACCATT 1620

GGAACAACCA AGCAGAATGT ACTTGATATT ATTTCAGGGT CCCATTGCTG TTCAGCCTTT 1680

GTTTTTACGT CATTACAAGC TGAGTAAAAT TCCTTCTGAT GATGTTATAA AAAAAAAAAA 1740

AAAAAAAA 1748

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 472 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Glu Ser Gln Phe Leu Arg Asp Thr Ile Cys Glu Glu Ser Leu Arg
1 5 10 15

Glu Lys Leu Gln Asp Gly Arg Ile Thr Ile Arg Glu Phe Phe Ile Leu
20 25 30

Leu Gln Val His Ile Leu Ile Gln Lys Pro Arg Gln Ser Asn Leu Pro
35 40 45

Gly Asn Phe Thr Val Asn Thr Pro Pro Thr Pro Glu Asp Leu Met Leu
50 55 60

Ser Gln Tyr Val Tyr Arg Pro Lys Ile Gln Ile Tyr Arg Glu Asp Cys
65 70 75 80

Glu Ala Arg Arg Gln Lys Ile Glu Glu Leu Lys Leu Ser Ala Ser Asn
85 90 95

Gln Asp Lys Leu Leu Val Asp Ile Asn Lys Asn Leu Trp Glu Lys Met
100 105 110

Arg His Cys Ser Asp Lys Glu Leu Lys Ala Phe Gly Ile Tyr Leu Asn
115 120 125

Lys Ile Lys Ser Cys Phe Thr Lys Met Thr Lys Val Phe Thr His Gln
130 135 140

Gly Lys Val Ala Leu Tyr Gly Lys Leu Val Gln Ser Ala Gln Asn Glu

```
145                 150                 155                 160

Arg Glu Lys Leu Gln Ile Lys Ile Asp Glu Met Asp Lys Ile Leu Lys
                165                 170                 175

Lys Ile Asp Asn Cys Leu Thr Glu Met Glu Thr Glu Thr Lys Asn Leu
            180                 185                 190

Glu Asp Glu Glu Lys Asn Asn Pro Val Glu Glu Trp Asp Ser Glu Met
        195                 200                 205

Arg Ala Ala Glu Lys Glu Leu Glu Gln Leu Lys Thr Glu Glu Glu Glu
    210                 215                 220

Leu Gln Arg Asn Leu Leu Glu Leu Glu Val Pro Lys Glu Gln Thr Leu
225                 230                 235                 240

Ala Gln Ile Asp Phe Met Gln Lys Gln Arg Asn Arg Thr Glu Glu Leu
                245                 250                 255

Leu Asp Gln Leu Ser Leu Ser Glu Trp Asp Val Val Glu Trp Ser Asp
            260                 265                 270

Asp Gln Ala Val Phe Thr Phe Val Tyr Asp Thr Ile Gln Leu Thr Ile
        275                 280                 285

Thr Phe Glu Glu Ser Val Val Gly Phe Pro Phe Leu Asp Lys Arg Tyr
    290                 295                 300

Arg Lys Ile Val Asp Val Asn Phe Gln Ser Leu Leu Asp Glu Asp Gln
305                 310                 315                 320

Ala Pro Pro Ser Ser Leu Leu Val His Lys Leu Ile Phe Gln Tyr Val
                325                 330                 335

Glu Glu Lys Glu Ser Trp Lys Lys Thr Cys Thr Thr Gln His Gln Leu
            340                 345                 350

Pro Lys Met Leu Glu Glu Phe Ser Leu Val Val His His Cys Arg Leu
        355                 360                 365

Leu Gly Glu Glu Ile Glu Tyr Leu Lys Arg Trp Gly Pro Asn Tyr Asn
    370                 375                 380

Leu Met Asn Ile Asp Ile Asn Asn Asn Glu Leu Arg Leu Leu Phe Ser
385                 390                 395                 400

Ser Ser Ala Ala Phe Ala Lys Phe Glu Ile Thr Leu Phe Leu Ser Ala
                405                 410                 415

Tyr Tyr Pro Ser Val Pro Leu Pro Ser Thr Ile Gln Asn His Val Gly
            420                 425                 430

Asn Thr Ser Gln Asp Asp Ile Ala Thr Ile Leu Ser Lys Val Pro Leu
        435                 440                 445
```

Glu Asn Asn Tyr Leu Lys Asn Val Val Lys Gln Ile Tyr Gln Asp Leu
450 455 460

Phe Gln Asp Cys His Phe Tyr His
465 470

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2298 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CTTTTCTTTG ATTGTCTCTG CTTTAGCGTC TCTAAATCCG GTCACCATGT CGGACCCCGA 60

AGGCGAGACC TTGCGAAGCA CCTTTCCCTC TTATATGGCC GAAGGCGAGC GGCTCTACCT 120

GTGCGGGGAA TTTTCTAAAG CCGCGCAGAG CTTCAGCAAC GCTCTTTACC TTCAGGATGG 180

AGACAAGAAC TGCCTGGTTG CTCGCTCAAA GTGCTTCCTG AAGATGGGAG ACTTGGAGAG 240

ATCCCTGAAG GATGCTGARG CTTCGCTCCA GAGTGACCCA GCTTTCTGTA AGGGGATTTT 300

GCAAAAGGCT GAGACACTGT ACACCATGGG AGACTTTGAG TTTGCCTTGG TATTCTATCA 360

TCGARGCTAC AAGCTGARGC CTGATCGGGA ATTCARARTT GGCATTCAGA AAGCCCAGGA 420

AGCCATCAAC AACTCAGTGG GAAGTCCTTC TTCCATTAAG CTGGAGAACA AAGGGGACCT 480

CTCCTTCTTA AGCAAGCAGG CTGAGAATAT AAAAGCCCAG CAGAAGCCTC AGCCCATGAA 540

ACACCTCTTA CACCCCACCA AGGGAGAGCC CAAGTGGAAG GCCTCGCTCA AGAGTGAGAA 600

GACTGTCCGC CAGCTTCTGG GGGAGCTCTA CGTGGACAAA GAGTATTTGG AGAAGCTCCT 660

ATTGGATGAA GACCTGATCA AAGGCACCAT GAAGGGCGGC CTGACTGTGG AGGACCTCAT 720

CATGACGGGC ATCAACTACC TGGATACTCA CAGCAACTTC TGGAGGCAGC AGAAGCCGAT 780

CTACGCCAGG GAGCGGGACC GGAAGCTGAT GCAAGAGAAA TGGCTGCGGG ACCACAAACG 840

CCGTCCCTCA CAGACAGCCC ATTACATCCT CAAGAGCCTG GAGGACATTG ATATGTTGCT 900

CACAAGTGGC AGTGCTGAAG GGAGTCTTCA GAAAGCTGAG AAAGTGCTGA AGAAGGTACT 960

GGAATGGAAC AAGGAAGAGG TACCCAACAA GGATGAACTG GTTGGAAACT TGTATAGCTG 1020

CATAGGGAAT GCCCAGATTG AGCTGGGGCA GATGGAGGCA GCCCTGCAGA GCCACAGAAA 1080

GGACYTGGAG ATCGCCAAGG AATATGACCT TCCTGATGCA AAATCGAGAG CCCTTGACAA 1140

CATTGGCAGA GTTTTTGCCA GAGTTGGGAA ATTCCAGCAA GCCATTGACA CGTGGGAAGA 1200

AAAGATCCCT CTGGCAAAAA CCACCCTGGA GAAGACCTGG CTGTTCCACG AGATCGGCCG 1260

CTGCTACTTG GAGCTGGACC AGGCCTGGCA GGCCCAGAAT TATGGCGAGA AGTCCCAGCA 1320

GTGTGCCGAG GAGGAAGGGG ACATTGAGTG GCAACTGAAT GCCAGTGTTC TGGTGGCCCA 1380

GGCACAAGTG AAGCTGAGAG ACTTCGAGTC AGCCGTGAAC AATTTTGAGA AGGCCCTGGA 1440

GAGAGCAAAG CTTGTGCATA ACAACGAGGC GCAGCAGGCC ATCATCAGTG CCTTGGACGA 1500

TGCCAACAAG GGTATCATCA GAGAACTGAG GAAAACCAAC TACGTGGAGA ATCTCAAAGA 1560

AAAAAGCGAG GGAGAAGCTT CACTGTATGA AGATAGAATA ATAACAAGAG AGAAGGACAT 1620

GAGGAGAGTG AGAGATGAGC CCGAGAAGGT GGTGAAGCAG TGGGACCATA GTGAGGATGA 1680

GAAAGAGACA GATGAGGACG ATGAGGCTTT TGGGGAAGCT CTGCAGAGCC CAGCAAGCGG 1740

AAAGCAGAGT GTGGAAGCAG GAAAAGCCAG AAGCGATTTG GGAGCAGTTG CCAAGGGCCT 1800

GTCAGGAGAA TTAGGCACAA GATCAGGAGA AACAGGCAGG AAGCTACTAG AAGCTGGCAG 1860

AAGAGAGTCA AGAGAAATTT ATAGGAGGCC TTCGGGAGAA TTAGAGCAAA GACTCTCAGG 1920

AGAATTCAGC AGACAGGAAC CAGAAGAACT AAAGAAACTT TCAGAAGTGG GCAGAAGAGA 1980

SCCAGAAGAA YTGGGAAAAA CACAATTTGG AGAAATAGGA GAAACGAAAA AAACAGGAAA 2040

TGAGATGGAA AAGGAATATG AATGAAGCCA TCGGTAGAGA TGAGGATCAG GAAGCTGGTG 2100

TTCAGAGGGA TCATGGGATT TTATTAAACT GGATTTTCAA GCGATTTGTC TGTTATAGGA 2160

AAAATGAGGG TTTTACTTYT GCTGCTTTCC ATCACTATTT TGCCATTAAA TAGGTGTCTT 2220

TCACTCTTGC MAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 2280

AAAAAAAAAA AAAAAAAA 2298

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 672 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Met Ser Asp Pro Glu Gly Glu Thr Leu Arg Ser Thr Phe Pro Ser Tyr
1               5                  10                  15

Met Ala Glu Gly Glu Arg Leu Tyr Leu Cys Gly Glu Phe Ser Lys Ala
            20                  25                  30

Ala Gln Ser Phe Ser Asn Ala Leu Tyr Leu Gln Asp Gly Asp Lys Asn
        35                  40                  45

Cys Leu Val Ala Arg Ser Lys Cys Phe Leu Lys Met Gly Asp Leu Glu
    50                  55                  60

Arg Ser Leu Lys Asp Ala Glu Ala Ser Leu Gln Ser Asp Pro Ala Phe
65                  70                  75                  80

Cys Lys Gly Ile Leu Gln Lys Ala Glu Thr Leu Tyr Thr Met Gly Asp
                85                  90                  95

Phe Glu Phe Ala Leu Val Phe Tyr His Arg Xaa Tyr Lys Leu Xaa Pro
            100                 105                 110

Asp Arg Glu Phe Xaa Xaa Gly Ile Gln Lys Ala Gln Glu Ala Ile Asn
        115                 120                 125

Asn Ser Val Gly Ser Pro Ser Ser Ile Lys Leu Glu Asn Lys Gly Asp
    130                 135                 140

Leu Ser Phe Leu Ser Lys Gln Ala Glu Asn Ile Lys Ala Gln Gln Lys
145                 150                 155                 160

Pro Gln Pro Met Lys His Leu Leu His Pro Thr Lys Gly Glu Pro Lys
                165                 170                 175

Trp Lys Ala Ser Leu Lys Ser Glu Lys Thr Val Arg Gln Leu Leu Gly
            180                 185                 190

Glu Leu Tyr Val Asp Lys Glu Tyr Leu Glu Lys Leu Leu Leu Asp Glu
        195                 200                 205

Asp Leu Ile Lys Gly Thr Met Lys Gly Gly Leu Thr Val Glu Asp Leu
    210                 215                 220

Ile Met Thr Gly Ile Asn Tyr Leu Asp Thr His Ser Asn Phe Trp Arg
225                 230                 235                 240

Gln Gln Lys Pro Ile Tyr Ala Arg Glu Arg Asp Arg Lys Leu Met Gln
                245                 250                 255

Glu Lys Trp Leu Arg Asp His Lys Arg Arg Pro Ser Gln Thr Ala His
            260                 265                 270

Tyr Ile Leu Lys Ser Leu Glu Asp Ile Asp Met Leu Leu Thr Ser Gly
```

275 280 285

Ser Ala Glu Gly Ser Leu Gln Lys Ala Glu Lys Val Leu Lys Lys Val
290 295 300

Leu Glu Trp Asn Lys Glu Glu Val Pro Asn Lys Asp Glu Leu Val Gly
305 310 315 320

Asn Leu Tyr Ser Cys Ile Gly Asn Ala Gln Ile Glu Leu Gly Gln Met
325 330 335

Glu Ala Ala Leu Gln Ser His Arg Lys Asp Leu Glu Ile Ala Lys Glu
340 345 350

Tyr Asp Leu Pro Asp Ala Lys Ser Arg Ala Leu Asp Asn Ile Gly Arg
355 360 365

Val Phe Ala Arg Val Gly Lys Phe Gln Gln Ala Ile Asp Thr Trp Glu
370 375 380

Glu Lys Ile Pro Leu Ala Lys Thr Thr Leu Glu Lys Thr Trp Leu Phe
385 390 395 400

His Glu Ile Gly Arg Cys Tyr Leu Glu Leu Asp Gln Ala Trp Gln Ala
405 410 415

Gln Asn Tyr Gly Glu Lys Ser Gln Gln Cys Ala Glu Glu Glu Gly Asp
420 425 430

Ile Glu Trp Gln Leu Asn Ala Ser Val Leu Val Ala Gln Ala Gln Val
435 440 445

Lys Leu Arg Asp Phe Glu Ser Ala Val Asn Asn Phe Glu Lys Ala Leu
450 455 460

Glu Arg Ala Lys Leu Val His Asn Asn Glu Ala Gln Gln Ala Ile Ile
465 470 475 480

Ser Ala Leu Asp Asp Ala Asn Lys Gly Ile Ile Arg Glu Leu Arg Lys
485 490 495

Thr Asn Tyr Val Glu Asn Leu Lys Glu Lys Ser Glu Gly Glu Ala Ser
500 505 510

Leu Tyr Glu Asp Arg Ile Ile Thr Arg Glu Lys Asp Met Arg Arg Val
515 520 525

Arg Asp Glu Pro Glu Lys Val Val Lys Gln Trp Asp His Ser Glu Asp
530 535 540

Glu Lys Glu Thr Asp Glu Asp Asp Glu Ala Phe Gly Glu Ala Leu Gln
545 550 555 560

Ser Pro Ala Ser Gly Lys Gln Ser Val Glu Ala Gly Lys Ala Arg Ser
565 570 575

Asp Leu Gly Ala Val Ala Lys Gly Leu Ser Gly Glu Leu Gly Thr Arg
580 585 590

Ser Gly Glu Thr Gly Arg Lys Leu Leu Glu Ala Gly Arg Arg Glu Ser
595 600 605

Arg Glu Ile Tyr Arg Arg Pro Ser Gly Glu Leu Glu Gln Arg Leu Ser
610 615 620

Gly Glu Phe Ser Arg Gln Glu Pro Glu Glu Leu Lys Lys Leu Ser Glu
625 630 635 640

Val Gly Arg Arg Xaa Pro Glu Glu Leu Gly Lys Thr Gln Phe Gly Glu
645 650 655

Ile Gly Glu Thr Lys Lys Thr Gly Asn Glu Met Glu Lys Glu Tyr Glu
660 665 670

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1010 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GGAAGAGCCA CCATCCCTGC CCCCGTTTTC CCACCGGGGA GTCTGTACAG AGATTTTTCT 60

ACGTTTTTAT TTTTTGCCTC AGAGGGATGG GATTGGGGAG GAGGGGATGG GCAGCGGAGG 120

GTTGGGGGCA TGGTCTGCAG GCTCATCTGT GTCCGCCTTT CACTCCACTA ATGCTGTCTC 180

AGTGTTTTCT CTCTCTCTCT TTCGAGCTTG CACTCCGGTA CCCGACCCGG CGCCCTGGCC 240

CATCCCATGC CGGGGGGCCA GTGGAAAGAA GACAGGCCGT CCAGCCCGTG CCCGCCTGCG 300

GCGGGGGCAC CCAGCAAGCC CGCCCACCGC CCGCTGCCTC ACCTGCTTCG CCACAGACTC 360

TTGTTCCCAG CCCCTTGGGG CCTCCGTGTT TGGGGTGGGG GAGCTGCTTA GAGACTGTGC 420

CCGTCCTCGG CCCCCCACCC TGAAGTGCCA GCACCACCAG CACCAGATCT TCCGCCGCCA 480

CACCGCATTG AGGACACGCC GGCCGGGCCG CTTCGTCTCA AGTTGTATAA AGTTGTCTCC 540

GTGTCCCCTC CTCCCTCTGC CCCCAGTGTT TCTTCTGATT TTTTTTTCCC CTTTCCCTCC 600

CTCCCTCTCC GCATTCTTCC CTTGGTTCAG CACAGGTAAA ACGGTTCCCC TCCCTCCCTG 660

CCTTCATGGA TCACCAGCTC ACGTCATGTT GCCTTCTCTT TTCTTTGTGT GTGTGTTTAT 720

TTAAGTTATT TTTCTTCCTC CTCTCCCTTT TCTTTTTGGC CCTCCCTCCC TCCCTCTTCT 780

GCCATGTAAC TGGAGGATGT GCTATGAGTT TGCAAACAGC TGGACTGTCA GGCTGCTTTT 840

TTTTCCAGAT GTTCTTCTTC TGCTTCCCCT TCCCCTCCTC TCCCCTCCTT TTCCTTCCTT 900

CCTTCCTTTC CTTGGAGCAC TGAGCACCAT TTGGAAGCTT GAGAGAAACC AAAATTAAAG 960

AGAGAAAGAG AGAGCGTGCA CGCTCCTGCT TTGTCAAAAA AAAAAAAAAA 1010

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 205 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Gly Ser Gly Gly Leu Gly Ala Trp Ser Ala Gly Ser Ser Val Ser
1 5 10 15

Ala Phe His Ser Thr Asn Ala Val Ser Val Phe Ser Leu Ser Leu Phe
20 25 30

Arg Ala Cys Thr Pro Val Pro Asp Pro Ala Pro Trp Pro Ile Pro Cys
35 40 45

Arg Gly Ala Ser Gly Lys Lys Thr Gly Arg Pro Ala Arg Ala Arg Leu
50 55 60

Arg Arg Gly His Pro Ala Ser Pro Pro Thr Ala Arg Cys Leu Thr Cys
65 70 75 80

Phe Ala Thr Asp Ser Cys Ser Gln Pro Leu Gly Ala Ser Val Phe Gly
85 90 95

Val Gly Glu Leu Leu Arg Asp Cys Ala Arg Pro Arg Pro Pro Thr Leu
100 105 110

Lys Cys Gln His His Gln His Gln Ile Phe Arg Arg His Thr Ala Leu
115 120 125

Arg Thr Arg Arg Pro Gly Arg Phe Val Ser Ser Cys Ile Lys Leu Ser
130 135 140

Pro Cys Pro Leu Leu Pro Leu Pro Pro Val Phe Leu Leu Ile Phe Phe

145 150 155 160

Ser Pro Phe Pro Pro Ser Leu Ser Ala Phe Phe Pro Trp Phe Ser Thr
165 170 175

Gly Lys Thr Val Pro Leu Pro Pro Cys Leu His Gly Ser Pro Ala His
180 185 190

Val Met Leu Pro Ser Leu Phe Phe Val Cys Val Phe Ile
195 200 205

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2409 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

ATTTYGCTCA TCAACCTCAT TATAGAACAT ATGATTTGTG ATACAGATCC TGAACTTGGA 60

GGAGCAGTCC AGCTTATGGG CCTGCTTCGA ACTTTAGTTG ACCCAGAGAA CATGCTAGCC 120

ACTGCCMATA AAACASAAAA GACTGAATTT CTGGGTTTCT TCTACAAGCA CTGTATGCAT 180

GTTCTCWCTG CTCCTTTACT AGCAAATACA ACAGAAGACA AACCTAGTAA AGATGATTTT 240

CAGACTGCCC AACTATTGGC ACTTGTATTG GAATTGTTAA CATTTTGTGT GGAGCACCAT 300

ACCTACCACA TAAAGAACTA CATTATTAAT AAGGATATCC TCCGGAGAGT GCTAGTTCTT 360

ATGGCCTCGA AGCATGCTTT CTTGGCATTA TGTGCCCTTC GTTTTAAAAG AAAGATTATT 420

GGATTAAAAG ATGAGTTTTA CAACCGCTAC ATAATGAAAA GTTTTTTGTT TGAACCAGTA 480

GTGAAAGCAT TTCTCAACAA TGGATCCCGC TACAATCTGA TGAACTCTGC CATAATAGAG 540

ATGTTTGAAT TTATTAGAGT GGAAGATATA AAATCATTAA CTGCTCATGT AATTGAAAAT 600

TACTGGAAAG CACTGGAAGA TGTAGATTAT GTACAGACAT TTAAAGGATT AAAACTGAGA 660

TTTGAACAAC AAAGAGAAAG GCAAGATAAT CCCAAACTTG ACAGTATGCG TTCCATTTTG 720

AGGAATCACA GATATCGAAG AGATGCCAGA ACACTAGAAG ATGAAGAAGA GATGTGGTTT 780

AACACAGATG AAGATGACAT GGAAGATGGA GAAGCTGTAG TGTCTCCATC TGACAAAACT 840

AAAAATGATG ATGATATTAT GGATCCAATA AGTAAATTCA TGGAAAGGAA GAAATTAAAA 900

GAAAGTGAGG AAAAGGAAGT GCTTCTGAAA ACAAACCTTT CTGGACGGCA GAGCCCAAGT 960

TTCAAGCTTT CCCTGTCCAG TGGAACGAAG ACTAACCTCA CCAGCCAGTC ATCTACAACA 1020

AATCTGCCTG GTTCTCCGGG ATCACCTGGA TCCCCAGGAT CTCCAGGCTC TCCTGGATCC 1080

GTACCTAAAA ATACATCTCA GACGGCAGCT ATTACTACAA AGGGAGGCCT CGTGGGTCTG 1140

GTAGATTATC CTGATGATGA TGAAGATGAT GATGAGGATG AAGATAAGGA AGATACGTTA 1200

CCATTGTCAA AGAAAGCAAA ATTTGATTCA TAATAATGGC AACGGCCTAG GATCAGTACC 1260

TGTTGAAAAA AACTGGTTCT CCACCCCTCC CCCATACAAA ATCCACAAAA AAGCGCAGTG 1320

GTCTCTTGTG AATGACTGAC ACAGATCAGC CTCTTACACT TGACTTCTGC TCATCAAGTG 1380

CCAATTCAAT GGAGCAGGAG GAGGGGATAT CATATATTTA GGGGAAAGAC TTAAGCCTTT 1440

GAGCTCTCCA GCTTGGACCA CACATTGCCC TTTTCTCAGG GAAGGAAATG GAAACAAAAA 1500

GCCAACAGGG CAGGGGTTTT GTAAGTGGAA CTCTGGATTG ACTGGTCAGT TGCTACAATC 1560

AGAATATGCT TTCTTGGACC ATGTTTGAGA CTCAGAAGAA TGGCCTTTCT GCCATAATTC 1620

TTCACTAGTC AAGAATGCCA GCAGTTTCTT TGTATAAAGA GACCTGCCTT TAAAATCATA 1680

CATTCTGAAC ATTTTAGTCA AGCTACAACA GGTTTGGAAA ACCTCTGTGG GGGAGGGGCG 1740

AGTATAAAGT TTTCCTCTTT TTTAACTGTT CCCTTTGCCC TTCAAACTGC AGATATTTTT 1800

TTTTTTAAGT GGGGACTTCT CCCTACTTGA TTAAAGATTG AGTGGAATTC TAGATGTGGT 1860

CATTTGTGTC ATAATTTTTT TGTTTTATTT TGTTTTTGAT TTTTTTTTTC CTCCCCTGAG 1920

TGTATGCTTA GTTGTTGAGT ATATATATTT GGGACCATTA AAACTTTTTT TGATGTAATA 1980

TAACCTAACG TTGTGCTGGT ACCTGTTTTA CCATGTGTAA TTTTTGTTCT ACATCACAGT 2040

TCTTAATTTG TTTAGAGTTT TATGAAAGAT GGTATAGTTT TTATTGACAA AAGCAAAGTA 2100

ATCTTACAAC TATGTGCATA CAAAAGCAAT ACTATTTTGT GACTAAATAT TTTATATTAA 2160

AATTTACATC AGCAACTGTC TTGAGAATTC AGGGAAATAG AATGGAATTT AAAACTTCAA 2220

CAGTTTTGTT AAATCTAGAA ACATGAAATT RGTATTCCAA AGAGATTCTG AAATTTCTTT 2280

TCTKGGGGAA ATGACGGTAC ATTAAATCAA AATTGRGGAT GGATGATTTA AAAACATTTG 2340

ACTTTTTTAAT AATAAAAAAGA AAAGTGAAGA GTAAGAGAAA TTGTAAAAAA AAAAAAAAAA 2400

AAAAAAAAA 2409

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 400 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```
Met Ile Cys Asp Thr Asp Pro Glu Leu Gly Gly Ala Val Gln Leu Met
1               5                  10                  15

Gly Leu Leu Arg Thr Leu Val Asp Pro Glu Asn Met Leu Ala Thr Ala
            20                  25                  30

Xaa Lys Thr Xaa Lys Thr Glu Phe Leu Gly Phe Phe Tyr Lys His Cys
        35                  40                  45

Met His Val Leu Xaa Ala Pro Leu Leu Ala Asn Thr Thr Glu Asp Lys
    50                  55                  60

Pro Ser Lys Asp Asp Phe Gln Thr Ala Gln Leu Leu Ala Leu Val Leu
65                  70                  75                  80

Glu Leu Leu Thr Phe Cys Val Glu His His Thr Tyr His Ile Lys Asn
                85                  90                  95

Tyr Ile Ile Asn Lys Asp Ile Leu Arg Arg Val Leu Val Leu Met Ala
            100                 105                 110

Ser Lys His Ala Phe Leu Ala Leu Cys Ala Leu Arg Phe Lys Arg Lys
        115                 120                 125

Ile Ile Gly Leu Lys Asp Glu Phe Tyr Asn Arg Tyr Ile Met Lys Ser
    130                 135                 140

Phe Leu Phe Glu Pro Val Val Lys Ala Phe Leu Asn Asn Gly Ser Arg
145                 150                 155                 160

Tyr Asn Leu Met Asn Ser Ala Ile Ile Glu Met Phe Glu Phe Ile Arg
                165                 170                 175

Val Glu Asp Ile Lys Ser Leu Thr Ala His Val Ile Glu Asn Tyr Trp
            180                 185                 190

Lys Ala Leu Glu Asp Val Asp Tyr Val Gln Thr Phe Lys Gly Leu Lys
        195                 200                 205

Leu Arg Phe Glu Gln Gln Arg Glu Arg Gln Asp Asn Pro Lys Leu Asp
    210                 215                 220

Ser Met Arg Ser Ile Leu Arg Asn His Arg Tyr Arg Arg Asp Ala Arg
```

225 230 235 240

Thr Leu Glu Asp Glu Glu Glu Met Trp Phe Asn Thr Asp Glu Asp Asp
245 250 255

Met Glu Asp Gly Glu Ala Val Val Ser Pro Ser Asp Lys Thr Lys Asn
260 265 270

Asp Asp Asp Ile Met Asp Pro Ile Ser Lys Phe Met Glu Arg Lys Lys
275 280 285

Leu Lys Glu Ser Glu Glu Lys Glu Val Leu Leu Lys Thr Asn Leu Ser
290 295 300

Gly Arg Gln Ser Pro Ser Phe Lys Leu Ser Leu Ser Ser Gly Thr Lys
305 310 315 320

Thr Asn Leu Thr Ser Gln Ser Ser Thr Thr Asn Leu Pro Gly Ser Pro
325 330 335

Gly Ser Pro Gly Ser Pro Gly Ser Pro Gly Ser Pro Gly Ser Val Pro
340 345 350

Lys Asn Thr Ser Gln Thr Ala Ala Ile Thr Thr Lys Gly Gly Leu Val
355 360 365

Gly Leu Val Asp Tyr Pro Asp Asp Asp Glu Asp Asp Asp Glu Asp Glu
370 375 380

Asp Lys Glu Asp Thr Leu Pro Leu Ser Lys Lys Ala Lys Phe Asp Ser
385 390 395 400

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 951 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GCCAGGCAGG GTGTGGGGGC AGCTGTGCCA ATCTACCTCA CAGGCCCACC CCCTGCCGGG 60

CATGCCGTGG GATCATGGGC AGGGAAGGCT CTGGGGGTCG GAGACACCGC TGCTTAGCAC 120

CCCCAGCCAG AACACCCTGA GGGTCTCGGG GCTCTGGAGA GAGTGGGGCG GGAGGAAGAA 180

TTGGCACCTT CCTAGGGAAG GAGACGAGCG CTTCGCCTTG ATTCTCCGAG AAGCCTCCGA 240

GAAGTGCTTT AAGTGTGTTT GCATGCSCCA GGCGGTGGGC AGCGGGGGCC TGTCCARCCC 300

TCTCCCGCCA TCCTTCCCCA AGTGACGTCC ACTGCCTTGT CACCAGCGAC CTGCCTGTCA 360

TGCCCACCCC CTGAGGAAGC ATGGGGACCC TAACACCCTG GTGCCCTGCA CCAGACAGGC 420

CGTGGTCAGG CCCAGGCCAC CGGCCGGGTT CTGCCACARC TTCCCACGTG CTTGCTGACA 480

TGCSTGTGCC TGTGTGTGGT GTCTGTTGCT GTGTCGTGAA ACTGTGACCA TCACTCAGTC 540

CAAACAAGTG AGTGGCCCTS GAGGCCACAG TTATGCAACT TTCAGTGTGT GTCATAACGA 600

CGTCACTGCT TTTTAAACTC GATAACTCTT TATTTTAGTA AAATGCCCAG GAGTCCTGGA 660

AGCTACGCGG ACTTGCAGAG GTTTTATTTT TTGGCCTTAG AATCTGCAGA AATTAGGAGG 720

CACCGAGCCC AGCGCAGCAG CCTCGGACCC GGATTGCGTT TGCCTTAGCG GATATGTTTA 780

TACAGATGAA TATAAAATGT TTTTTTCTTT GGGCTTTTTG CTTCTTTTTT CCCCCCCTTC 840

TCACCTTCCC TTCTCCCTGA CCCCACCCCC CAAAAAAGCT ACTTCTTCAT TCCGTGGTAC 900

GATTATTTTT TTTAACTAAA GGAAGATAAA ATTCTAAAAA AAAAAAAAAA A 951

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 87 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Pro Trp Asp His Gly Gln Gly Arg Leu Trp Gly Ser Glu Thr Pro
1 5 10 15

Leu Leu Ser Thr Pro Ser Gln Asn Thr Leu Arg Val Ser Gly Leu Trp
20 25 30

Arg Glu Trp Gly Gly Arg Lys Asn Trp His Leu Pro Arg Glu Gly Asp
35 40 45

Glu Arg Phe Ala Leu Ile Leu Arg Glu Ala Ser Glu Lys Cys Phe Lys
50 55 60

Cys Val Cys Met Xaa Gln Ala Val Gly Ser Gly Gly Leu Ser Xaa Pro
65 70 75 80

Leu Pro Pro Ser Phe Pro Lys

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1899 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGCCGCTTGT GTCCACGGGA CGCGGGCGGA TCTTCTCCGG CCATGAGGAA GCCAGCCGCT 60

GGCTTCCTTC CCTCACTCCT GAAGGTGCTG CTCCTGCCTC TGGCACCTGC CGCAGCCCAG 120

GATTCGACTC AGGCCTCCAC TCCAGGCAGC CCTCTCTCTC CTACCGAATA CGAACGCTTC 180

TTCGCACTGC TGACTCCAAC CTGGAAGGCA GAGACTACCT GCCGTCTCCG TGCAACCCAC 240

GGCTGCCGGA ATCCCACACT CGTCCAGCTG GACCAATATG AAAACCACGG CTTAGTGCCC 300

GATGGTGCTG TCTGCTCCAA CCTCCCTTAT GCCTCCTGGT TTGAGTCTTT CTGCCAGTTC 360

ACTCACTACC GTTGCTCCAA CCACGTCTAC TATGCCAAGA GAGTCCTGTG TTCCCAGCCA 420

GTCTCTATTC TCTCACCTAA CACTCTCAAG GAGATAGAAG CTTCAGCTGA AGTCTCACCC 480

ACCACGATGA CCTCCCCCAT CTCACCCCAC TTCACAGTGA CAGAACGCCA GACCTTCCAG 540

CCCTGGCCTG AGAGGCTCAG CAACAACGTG GAAGAGCTCC TACAATCCTC CTTGTCCCTG 600

GGAGGCCAGG AGCAAGCGCC AGAGCACAAG CAGGAGCAAG GAGTGGAGCA CAGGCAGGAG 660

CCGACACAAG AACACAAGCA GGAAGAGGGG CAGAAACAGG AAGAGCAAGA AGAGGAACAG 720

GAAGAGGAGG GAAAGCAGGA AGAAGGACAG GGGACTAAGG AGGGACGGGA GGCTGTGTCT 780

CAGCTGCAGA CAGACTCAGA GCCCAAGTTT CACTCTGAAT CTCTATCTTC TAACCCTTCC 840

TCTTTTGCTC CCCGGGTACG AGAAGTAGAG TCTACTCCTA TGATAATGGA GAACATCCAG 900

GAGCTCATTC GATCAGCCCA GGAAATAGAT GAAATGAATG AAATATATGA TGAGAACTCC 960

TACTGGAGAA ACCAAAACCC TGGCAGCCTC CTGCAGCTGC CCCACACAGA GGCCTTGCTG 1020

GTGCTGTGCT ATTCGATCGT GGAGAATACC TGCATCATAA CCCCCACAGC CAAGGCCTGG 1080

AAGTACATGG AGGAGGAGAT CCTTGGTTTC GGGAAGTCGG TCTGTGACAG CCTTGGGCGG 1140

CGACACATGT CTACCTGTGC CCTCTGTGAC TTCTGCTCCT TGAAGCTGGA GCAGTGCCAC 1200

TCAGAGGCCA GCCTGCAGCG GCAACAATGC GACACCTCCC ACAAGACTCC CTTTGTCAGC 1260

CCCTTGCTTG CCTCCCAGAG CCTGTCCATC GGCAACCAGG TAGGGTCCCC AGAATCAGGC 1320

CGCTTTTACG GGCTGGATTT GTACGGTGGG CTCCACATGG ACTTCTGGTG TGCCCGGCTT 1380

GCCACGAAAG GCTGTGAAGA TGTCCGAGTC TCTGGGTGGC TCCAGACTGA GTTCCTTAGC 1440

TTCCAGGATG GGGATTTCCC TACCAAGATT TGTGACACAG ACTATATCCA GTACCCAAAC 1500

TACTGTTCCT TCAAAAGCCA GCAGTGTCTG ATGAGAAACC GCAATCGGAA GGTGTCCCGC 1560

ATGAGATGTC TGCAGAATGA GACTTACAGT GCGCTGAGCC TGGCAAAAGT GAGGACGTTG 1620

TGCTTTCGAT GGAGCCAGGA GTTCAGCACC TTGACTCTAG GCCAGTTCGG ATGAGCTKGS 1680

GTTTATTTTG CCCACACCCC AGCCCAACCT GCCCASGTTC TCTATTGTTT TGAGACCCCA 1740

TTGCTTTCAG GCTGCCCCTT CTGGGTCTGT TACTCGGCCC CTAMTCACAT TTCCTTGGGT 1800

TGGAGCAACA GTCCCAGAGA GGGCCACGGT GGGAGCTGCG CCCTCCTTAA AAGATGACTT 1860

TACATAAAAT GTTGATCTTC AAAAAAAAAA AAAAAAAAA 1899

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 543 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met Arg Lys Pro Ala Ala Gly Phe Leu Pro Ser Leu Leu Lys Val Leu
1 5 10 15

Leu Leu Pro Leu Ala Pro Ala Ala Ala Gln Asp Ser Thr Gln Ala Ser
20 25 30

Thr Pro Gly Ser Pro Leu Ser Pro Thr Glu Tyr Glu Arg Phe Phe Ala
35 40 45

Leu Leu Thr Pro Thr Trp Lys Ala Glu Thr Thr Cys Arg Leu Arg Ala
50 55 60

Thr His Gly Cys Arg Asn Pro Thr Leu Val Gln Leu Asp Gln Tyr Glu
65 70 75 80

```
Asn His Gly Leu Val Pro Asp Gly Ala Val Cys Ser Asn Leu Pro Tyr
                85                  90                  95

Ala Ser Trp Phe Glu Ser Phe Cys Gln Phe Thr His Tyr Arg Cys Ser
            100                 105                 110

Asn His Val Tyr Tyr Ala Lys Arg Val Leu Cys Ser Gln Pro Val Ser
        115                 120                 125

Ile Leu Ser Pro Asn Thr Leu Lys Glu Ile Glu Ala Ser Ala Glu Val
    130                 135                 140

Ser Pro Thr Thr Met Thr Ser Pro Ile Ser Pro His Phe Thr Val Thr
145                 150                 155                 160

Glu Arg Gln Thr Phe Gln Pro Trp Pro Glu Arg Leu Ser Asn Asn Val
                165                 170                 175

Glu Glu Leu Leu Gln Ser Ser Leu Ser Leu Gly Gly Gln Glu Gln Ala
            180                 185                 190

Pro Glu His Lys Gln Glu Gln Gly Val Glu His Arg Gln Glu Pro Thr
        195                 200                 205

Gln Glu His Lys Gln Glu Glu Gly Gln Lys Gln Glu Glu Gln Glu Glu
    210                 215                 220

Glu Gln Glu Glu Glu Gly Lys Gln Glu Glu Gly Gln Gly Thr Lys Glu
225                 230                 235                 240

Gly Arg Glu Ala Val Ser Gln Leu Gln Thr Asp Ser Glu Pro Lys Phe
                245                 250                 255

His Ser Glu Ser Leu Ser Ser Asn Pro Ser Ser Phe Ala Pro Arg Val
            260                 265                 270

Arg Glu Val Glu Ser Thr Pro Met Ile Met Glu Asn Ile Gln Glu Leu
        275                 280                 285

Ile Arg Ser Ala Gln Glu Ile Asp Glu Met Asn Glu Ile Tyr Asp Glu
    290                 295                 300

Asn Ser Tyr Trp Arg Asn Gln Asn Pro Gly Ser Leu Leu Gln Leu Pro
305                 310                 315                 320

His Thr Glu Ala Leu Leu Val Leu Cys Tyr Ser Ile Val Glu Asn Thr
                325                 330                 335

Cys Ile Ile Thr Pro Thr Ala Lys Ala Trp Lys Tyr Met Glu Glu Glu
            340                 345                 350

Ile Leu Gly Phe Gly Lys Ser Val Cys Asp Ser Leu Gly Arg Arg His
        355                 360                 365

Met Ser Thr Cys Ala Leu Cys Asp Phe Cys Ser Leu Lys Leu Glu Gln
```

370 375 380

Cys His Ser Glu Ala Ser Leu Gln Arg Gln Gln Cys Asp Thr Ser His
385 390 395 400

Lys Thr Pro Phe Val Ser Pro Leu Leu Ala Ser Gln Ser Leu Ser Ile
405 410 415

Gly Asn Gln Val Gly Ser Pro Glu Ser Gly Arg Phe Tyr Gly Leu Asp
420 425 430

Leu Tyr Gly Gly Leu His Met Asp Phe Trp Cys Ala Arg Leu Ala Thr
435 440 445

Lys Gly Cys Glu Asp Val Arg Val Ser Gly Trp Leu Gln Thr Glu Phe
450 455 460

Leu Ser Phe Gln Asp Gly Asp Phe Pro Thr Lys Ile Cys Asp Thr Asp
465 470 475 480

Tyr Ile Gln Tyr Pro Asn Tyr Cys Ser Phe Lys Ser Gln Gln Cys Leu
485 490 495

Met Arg Asn Arg Asn Arg Lys Val Ser Arg Met Arg Cys Leu Gln Asn
500 505 510

Glu Thr Tyr Ser Ala Leu Ser Leu Ala Lys Val Arg Thr Leu Cys Phe
515 520 525

Arg Trp Ser Gln Glu Phe Ser Thr Leu Thr Leu Gly Gln Phe Gly
530 535 540

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 722 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CGACCTTCCC AGCAATATGC ATCTTGCACG TCTGGTCGGC TCCTGCTCCC TCCTTCTGCT 60

ACTGGGGGCC CTGTCTGGAT GGGCGGCCAG CGATGACCCC ATTGAGAAGG TCATTGAAGG 120

GATCAACCGA GGGCTGAGCA ATGCAGAGAG AGAGGTGGGC AAGGCCCTGG ATGGCATCAA 180

CAGTGGAATC ACGCATGCCG GAAGGGAAGT GGAGAAGGTT TTCAACGGAC TTAGCAACAT 240

GGGGAGCCAC ACCGGCAAGG AGTTGGACAA AGGCGTCCAG GGGCTCAACC ACGGCATGGA 300

CAAGGTTGCC CATGAGATCA ACCATGGTAT TGGACAAGCA GGAAAGGAAG CAGAGAAGCT 360

TGGCCATGGG GTCAACAACG CTGCTGGACA GGGCAACCAT CAAAGCGGAT CTTCCAGCCA 420

TCAAGGAGGG GCCACAACCA CGCCGTTAGC CTCTGGGGCC TCGGTCAACA CGCCTTTCAT 480

CAACCTTCCC GCCCTGTGGA GGAGCGTCGC CAACATCATG CCCTAAACTG GCATCCGGCC 540

TTGCTGGGAG AATAATGTCG CCGTTGTCAC ATCAGCTGAC ATGACCTGGA GGGGTTGGGG 600

GTGGGGGACA GGTTTCTGAA ATCCCTGAAG GGGGTTGTAC TGGGATTTGT GAATAAACTT 660

GATACACTAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 720

AA 722

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 169 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met His Leu Ala Arg Leu Val Gly Ser Cys Ser Leu Leu Leu Leu Leu
1 5 10 15

Gly Ala Leu Ser Gly Trp Ala Ala Ser Asp Asp Pro Ile Glu Lys Val
20 25 30

Ile Glu Gly Ile Asn Arg Gly Leu Ser Asn Ala Glu Arg Glu Val Gly
35 40 45

Lys Ala Leu Asp Gly Ile Asn Ser Gly Ile Thr His Ala Gly Arg Glu
50 55 60

Val Glu Lys Val Phe Asn Gly Leu Ser Asn Met Gly Ser His Thr Gly
65 70 75 80

Lys Glu Leu Asp Lys Gly Val Gln Gly Leu Asn His Gly Met Asp Lys
85 90 95

Val Ala His Glu Ile Asn His Gly Ile Gly Gln Ala Gly Lys Glu Ala
100 105 110

Glu Lys Leu Gly His Gly Val Asn Asn Ala Ala Gly Gln Gly Asn His

115 120 125

Gln Ser Gly Ser Ser Ser His Gln Gly Gly Ala Thr Thr Thr Pro Leu
130 135 140

Ala Ser Gly Ala Ser Val Asn Thr Pro Phe Ile Asn Leu Pro Ala Leu
145 150 155 160

Trp Arg Ser Val Ala Asn Ile Met Pro
165

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1240 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

AATGGCTTTT CTTCCTTCCT GGGTTTGTGT ACTAGTTGGT TCCTTTTCTG CTTCCTTAGC 60

AGGGACTTCC AATCTCTCAG AGACAGAGCC CCCTCTGTGG AAGGAGAGTC CTGGTCAGCT 120

CAGTGACTAC AGGGTGGAGA ACAGCATGTA CATTATTAAT CCCTGGGTAT ACCTTGAGAG 180

AATGGGGATG TATAAAATCA TATTGAATCA GACAGCCAGG TATTTTGCAA AATTTGCACC 240

AGATAATGAA CAGAATATTT TATGGGGGTT GCCTCTGCAG TATGGCTGGC AATATAGGAC 300

AGGCAGATTA GCTGATCCAA CCCGAAGGAC AAACTGTGGC TATGAATCTG GAGATCATAT 360

GTGCATCTCT GTGGACAGTT GGTGGGCTGA TTTGAATTAT TTTCTGTCTT CATTACCCTT 420

TCTTGCTGCG GTTGATTCTG GTGTAATGGG GATATCATCA GACCAAGTCA GGCTTTTGCC 480

CCCACCCAAG AATGAGAGGA AGTTTTGTTA TGATGTTTCT AGCTGTCGTT CATCCTTCCC 540

TGAGACAATG AACAAGTGGA ACACCTTTTA CCAGTATTTG CAGTCACCTT TTAGTAAGTT 600

TGATGATCTG TTGAAGTACT TATGGGCTGC ACACACTTCA ACCTTGGCAG ATAATATCAA 660

AAGTTTTGAA GACAGATATG ATTATTATTC TAAAGCAGAA GCGCATTTTG AGAGAAGTTG 720

GGTACTGGCT GTGGATCATT TAGCTGCAGT CCTCTTTCCT ACAACCTTGA TTAGATCATA 780

TAAGTTCCAG AAGGGCATGC CACCACGAAT TCTTCTTAAT ACTGATGTAG CCCCTTTCAT 840

CAGTGACTTT ACTGCTTTTC AGAATGTAGT CCTGGTTCTT CTAAATATGC TTGACAATGT 900

GGATAAATCT ATAGGTTATC TTTGTACAGA AAAATCTAAT GTATATAGAG ATCATTCGGA 960

ATCTAGCTCT AGAAGTTATG GAAATAACTC CTGAAACATT TAACTTCAAA CTTCAGGAAA 1020

TGATTAATGA ATTAAAAATG AAAAACTCGA ACTTGACAAT CAGTAATTTC AAAAAATTAA 1080

TGTCATCATG ACCATGTAGT TTATTCTTTC TGATATTTTT GATTTATGCT TATTTGTTAA 1140

GATCTTGTAC ATGTATTAAA AACTTAAATT AAATGCATTC AAGTTAAAAA AAAAAAAAAA 1200

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 1240

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 330 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Met Ala Phe Leu Pro Ser Trp Val Cys Val Leu Val Gly Ser Phe Ser
1 5 10 15

Ala Ser Leu Ala Gly Thr Ser Asn Leu Ser Glu Thr Glu Pro Pro Leu
20 25 30

Trp Lys Glu Ser Pro Gly Gln Leu Ser Asp Tyr Arg Val Glu Asn Ser
35 40 45

Met Tyr Ile Ile Asn Pro Trp Val Tyr Leu Glu Arg Met Gly Met Tyr
50 55 60

Lys Ile Ile Leu Asn Gln Thr Ala Arg Tyr Phe Ala Lys Phe Ala Pro
65 70 75 80

Asp Asn Glu Gln Asn Ile Leu Trp Gly Leu Pro Leu Gln Tyr Gly Trp
85 90 95

Gln Tyr Arg Thr Gly Arg Leu Ala Asp Pro Thr Arg Arg Thr Asn Cys
100 105 110

Gly Tyr Glu Ser Gly Asp His Met Cys Ile Ser Val Asp Ser Trp Trp
115 120 125

Ala Asp Leu Asn Tyr Phe Leu Ser Ser Leu Pro Phe Leu Ala Ala Val
130 135 140

Asp Ser Gly Val Met Gly Ile Ser Ser Asp Gln Val Arg Leu Leu Pro

145 150 155 160

Pro Pro Lys Asn Glu Arg Lys Phe Cys Tyr Asp Val Ser Ser Cys Arg
165 170 175

Ser Ser Phe Pro Glu Thr Met Asn Lys Trp Asn Thr Phe Tyr Gln Tyr
180 185 190

Leu Gln Ser Pro Phe Ser Lys Phe Asp Asp Leu Leu Lys Tyr Leu Trp
195 200 205

Ala Ala His Thr Ser Thr Leu Ala Asp Asn Ile Lys Ser Phe Glu Asp
210 215 220

Arg Tyr Asp Tyr Tyr Ser Lys Ala Glu Ala His Phe Glu Arg Ser Trp
225 230 235 240

Val Leu Ala Val Asp His Leu Ala Ala Val Leu Phe Pro Thr Thr Leu
245 250 255

Ile Arg Ser Tyr Lys Phe Gln Lys Gly Met Pro Pro Arg Ile Leu Leu
260 265 270

Asn Thr Asp Val Ala Pro Phe Ile Ser Asp Phe Thr Ala Phe Gln Asn
275 280 285

Val Val Leu Val Leu Leu Asn Met Leu Asp Asn Val Asp Lys Ser Ile
290 295 300

Gly Tyr Leu Cys Thr Glu Lys Ser Asn Val Tyr Arg Asp His Ser Glu
305 310 315 320

Ser Ser Ser Arg Ser Tyr Gly Asn Asn Ser
325 330

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2261 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GCAGCACCAG CCGTCTGCAG CTCCGGCCGC CACTTGCGCC TCTCCAGCCT CCGCAGGCCC 60

AACCGCCGCC AGCACCATGG CCAGCACCAT TTCCGCCTAC AAGGAGAAGA TGAAGGAGCT 120

GTCGGTGCTG TCGCTCATCT GCTCCTGCTT CTACACACAG CCGCACCCCA ATACCGTCTA 180

CCAGTACGGG GACATGGAGG TGAAGCAGCT GGACAAGCGG GCCTCAGGCC AGAGCTTCGA 240

GGTCATCCTC AAGTCCCCTT CTGACCTGTC CCCAGAGAGC CCTATGCTCT CCTCCCCACC 300

CAAGAAGAAG GACACCTCCC TGGAGGAGCT GCAAAAGCGG CTGGAGGCAG CCGAGGAGCG 360

GAGGAAGACG CAGGAGGCGC AGGTGCTGAA GCAGCTGGCG GAGCGGCGCG AGCACGAGCG 420

CGAGGTGCTG CACAAGGCGC TGGAGGAGAA TAACAACTTC AGCCGCCAGG CGGAGGAGAA 480

GCTCAACTAC AAGATGGAGC TCAGCAAGGA GATCCGCGAG GCACACCTGG CCGCACTGCG 540

CGAGCGGCTG CGCGAGAAGG AGCTGCACGC GGCCGAGGTG CGCAGGAACA AGGAGCAGCG 600

AGAAGAGATG TCGGGCTAAG GGCCCGGGAC GGGCGGCGCC CATCCTGCGA CAGAACACGT 660

TCGGGTTTTG GTTTTGTTTC GTTCACCTCT GTCTAGATGC AACTTTTGTT CCTCCTCCCC 720

CACCCCAGCC CCCAGCTTCA TGCTTCTCTT CCGCACTCAG CCGCCCTGCC CTGTCCTCGT 780

GGTGAGTCGC TGACCACGGC TTCCCCTGCA GGAGCCGCCG GGCGTGAGAC GCGGTCCCTC 840

GGTGCAGACA CCAGGCCGGG CGCGGCTGGG TCCCCCGGGG GCCCTGTGAG AGAGGTGGCG 900

GTGACCGTGG TAAACCCAGG GCGGTGGCGT GGGATCGCGG GTCCTTACGC TGGGCTGTCT 960

GGTCAGCACG TGCAGGTCAG GGCAGGTCCT CTGAGCCGGC GCCCCTGGCC AGCAGGCGAG 1020

GCTACAGTAC CTGCTGTCTT TCCAGGGGGA AGGGGCTCCC CATGAGGGAG GGGCGACGGG 1080

GGAGGGGGGT GATGGTGCCT GGGAGCCTGC GTGTGCAGCC GGTGCTTGTT GAACTGGCAG 1140

GCGGGTGGGT GGGGGCTGCA GCTTTCCTTA ATGTGGTTGC ACAGGGGTCC TCTGAGACCA 1200

CCTGGCGTGA GGTGGACACC CTGGGCCTTC CTGGAAGCCT GCAGTTGGGG GCCTGCCCTG 1260

AGTCTGCTGG GGAGTGGGCA TTCTCTGCCA GGGACCCATG AGCAGGCTGC ATGGTCTAGA 1320

GGTTGTGGGC AGCATGGACA GTCCCCCACT CAGAAGTGCA AGAGTTCCAA AGAGCCTCTG 1380

GCCCAGGCCC CTCCCCACCA GGGCTTTGCA GATGTCCTTG AAAGACCCAC CCTAGAGCCC 1440

TTTGGAGTGC TGGCCCCTCC TGTGCCCTCT GCCCTGGTGG AAGCGGCAGC CACAAGTCCT 1500

CCTCAGGGAG CCCCAAGGGG GATTTTGTGG GACCGCTGCC CACAGATCCA GGTGTTGGAA 1560

GGGCAGCGGG TAAGGTTCCC AAGCCAGCCC CAACACCCTT CCCACTTGGC ACCCAGAGGG 1620

GGCTGTGGGT GGAGGCCTGA CTCCAGGCCT CTCCTGCCCA CACCCTCTGG GCTGAGTTCC 1680

TTCTTTCCCT TGGACGCCCA GTGCTGGCCT TGGAGGACGG TCAGCTGGAG GATGGCGGTG 1740

GGGGAGGCTG TCTTTGTACC ACTGCAGCAT CCCCCACTTC TCCACGGAAG CCCCATCCCA 1800

AAGCTGCTGC CTGGCCCCTT GCTGTAAAGT GTGAAGGGGG CGGCTGAGTT CTCTTAGGAC 1860

CCAGAGCCAG GGCCCTCAAC TTCCATCCTG CGGGAGGCCT TGGCCGGGCA CTGCCAGTGT 1920

CTTCCAGAGC CACACCCAGG GACCACGGGA GGATCCTGAC CCCTGCAGGG CTCAGGGGTC 1980

AGCAGGGACC CACTGCCCCA TCTCCCTCTC CCCACCAAGA CAGCCCCAGA AGGAGCAGCC 2040

AGCTGGGATG GGAACCCAAG GCTGTCCACA TCTGGCTTTT GTGGGACTCA GAAAGGGAAG 2100

CAGAACTGAG GGCTGGGATA TTCCTCATGG TGGCAGCGCT CATAGCGAAA GCCTACTGTA 2160

ATATGCACCC ATCTCATCCA CGTAGTAAAG TGAACTTAAA AATTCAATCA AATGAACAAT 2220

TAAATAAACA CCTGTGTGTT TAAGAAAAAA AAAAAAAAAA A 2261

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 180 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met Ala Ser Thr Ile Ser Ala Tyr Lys Glu Lys Met Lys Glu Leu Ser
1 5 10 15

Val Leu Ser Leu Ile Cys Ser Cys Phe Tyr Thr Gln Pro His Pro Asn
20 25 30

Thr Val Tyr Gln Tyr Gly Asp Met Glu Val Lys Gln Leu Asp Lys Arg
35 40 45

Ala Ser Gly Gln Ser Phe Glu Val Ile Leu Lys Ser Pro Ser Asp Leu
50 55 60

Ser Pro Glu Ser Pro Met Leu Ser Ser Pro Pro Lys Lys Lys Asp Thr
65 70 75 80

Ser Leu Glu Glu Leu Gln Lys Arg Leu Glu Ala Ala Glu Glu Arg Arg
85 90 95

Lys Thr Gln Glu Ala Gln Val Leu Lys Gln Leu Ala Glu Arg Arg Glu
100 105 110

His Glu Arg Glu Val Leu His Lys Ala Leu Glu Glu Asn Asn Asn Phe
115 120 125

Ser Arg Gln Ala Glu Glu Lys Leu Asn Tyr Lys Met Glu Leu Ser Lys
130 135 140

Glu Ile Arg Glu Ala His Leu Ala Ala Leu Arg Glu Arg Leu Arg Glu
145 150 155 160

Lys Glu Leu His Ala Ala Glu Val Arg Arg Asn Lys Glu Gln Arg Glu
165 170 175

Glu Met Ser Gly
180

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3109 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GGCCAAAGAG GCCTAGGAGC CTCGTGGCTG CGTCACCGCC GCCCCCCCAG ACAAGATGGA 60

CACCGCGGAG GAAGACATAT GTAGAGTGTG TCGGTCAGAA GGAACACCTG AGAAACCGCT 120

TTATCATCCT TGTGTATGTA CTGGCAGTAT TAAGTTTATC CATCAAGAAT GCTTAGTTCA 180

ATGGCTGAAA CACAGTCGAA AAGAATACTG TGAATTATGC AAGCACAGAT TTGCTTTTAC 240

ACCAATTTAT TCTCCAGATA TGCCTTCACG GCTTCCAATT CAAGACATAT TTGCTGGACT 300

GGTTACAAGT ATTGGCACTG CAATACGATA TTGGTTTCAT TATACACTTG TGGCCTTTGC 360

ATGGTTGGGA GTTGTTCCTC TTACAGCATG CCGCATCTAC AAGTGCTTGT TTACTGGCTC 420

CCGTGAGCTC ACTACTGACG CTGCCCATTA GATATGCTGT CAACCGGAAA ATTTGTTGGC 480

AGATTGTTTG CAGGGTTGTT TTGTGGTGAC GTGCACACTG TGTGCATTCA TCAGCCTGGT 540

GTGGTTGAGA GAGCAGATAG TCCATGGGGG AGCACCAATT TGGTTGGAGC ATGCTGCCCC 600

ACCGTTCAAT GCTGCGGGGC ATCACCAAAA TGAGGCTCCA GCAGGAGGAA ATGGTGCAGA 660

AAATGTTGCT GCTGATCAGC CTGCTAACCC ACCAGCTGAG AACGCAGTGG TGGGGGAAAA 720

CCCTGATGCC CAGGATGACC AGGCAGAAGA GGAGGAGGAG GACAATGAGG AGGAAGATGA 780

CGCTGGTGTG GAGGATGGCG GCAGATGCTA ATAACGGAGC CCAGGATGAC ATGAATTGGA 840

ATGCTTTAGA ATGGGACCGA GCTGCTGAAG AGCTTACATG GGAAAGAATG CTAGGACTTG 900

ATGGATCACT AGTTTTTCTG GAACATGTCT TCTGGGTGGT ATCTTTAAAT ACACTGTTCA 960

TTCTTGTTTT TGCATTTTGC CCTTACCATA TTGGTCATTT CTCCCTTGTT GGTTTGGGAT 1020

TTGAAGAACA CGTCCAAGCA TCTCATTTTG AAGGCCTAAT CACAACCATA GTTGGGTATA 1080

TACTTTTAGC AATAACACTG ATAATTTGTC ATGGCTTGGC AACTCTTGTG AAATTTCATA 1140

GATCTCGTCG CTTACTGGGA GTCTGCTATA TTGTTGTTAA GGTCTCTTTG TTAGTGGTGG 1200

TAGAAATTGG AGTATTCCCT CTCATTTGTG GTTGGTGGCT GGATATCTGT TCCTTGGAAA 1260

TGTTTGATGC TACTCTGAAA GATCGAGAAC TGAGCTTTCA GTCGGCTCCA GGTACTACCA 1320

TGTTTCTGCA TTGGCTAGTG GGAATGGTAT ATGTCTTCTA CTTTGCCTCC TTCATTCTAT 1380

TACTGAGAGA GGTACTTCGA CCTGGTGTCC TGTGGTTTCT AAGGAATTTG AATGATCCAG 1440

ATTTCAATCC AGTACAGGAA ATGATCCATT TGCCAATATA TAGGCATCTC CGAAGATTTA 1500

TTTTGTCAGT GATTGTCTTT GGCTCCATTG TCCTCCTGAT GCTTTGGCTT CCTATACGTA 1560

TAATTAAGAG TGTGCTGCCT AATTTTCTTC CATACAATGT CATGCTCTAC AGTGATGCTC 1620

CAGTGAGTGA ACTGTCCCTC GAGCTGCTTC TGCTTCAGGT TGTCTTGCCA GCATTACTCG 1680

AACAGGGACA CACGAGGCAG TGGCTGAAGG GGCTGGTGCG AGCGTGGACT GTGACCGCCG 1740

GATACTTGCT GGATCTTCAT TCTTATTTAT TGGGAGACCA GGAAGAAAAT GAAAACAGTG 1800

CAAATCAACA AGTTAACAAT AATCAGCATG CTCGAAATAA CAACGCTATT CCTGTGGTGG 1860

GAGAAGGCCT TCATGCAGCC CACCAAGCCA TACTCCAGCA GGGAGGGCCT GTTGGCTTTC 1920

AGCCTTACCG CCGACCTTTA AATTTTCCAC TCAGGATATT TCTGTTGATT GTCTTCATGT 1980

GTATAACATT ACTGATTGCC AGCCTCATCT GCCTTACTTT ACCAGTATTT GCTGGCCGTT 2040

GGTTAATGTC GTTTTGGACG GGGACTGCCA AAATCCATGA GCTCTACACA GCTGCTTGTG 2100

GTCTCTATGT TTGCTGGCTA ACCATAAGGG CTGTGACGGT GATGGTGGCA TGGATGCCTC 2160

AGGGACGCAG AGTGATCTTC CAGAAGGTTA AAGAGTGGTC TCTCATGATC ATGAAGACTT 2220

TGATAGTTGC GGTGCTGTTG GCTGGAGTTG TCCCTCTCCT TCTGGGGCTC CTGTTTGAGC 2280

TGGTCATTGT GGCTCCCCTG AGGGTTCCCT TGGATCAGAC TCCTCTTTTT TATCCATGGC 2340

AGGACTGGGC ACTTGGAGTC CTGCATGCCA AAATCATTGC AGCTATAACA TTGATGGGTC 2400

CTCAGTGGTG GTTGAAAACT GTAATTGAAC AGGTTTACGC AAATGGCATC CGGAACATTG 2460

ACCTTCACTA TATTGTTCGT AAACTGGCAG CTCCCGTGAT CTCTGTGCTG TTGCTTTCCC 2520

TGTGTGTACC TTATGTCATA GCTTCTGGTG TTGTTCCTTT ACTAGGTGTT ACTGCGGAAA 2580

TGCAAAACTT AGTCCATCGG CGGATTTATC CATTTTTACT GATGGTCGTG GTATTGATGG 2640

CAATTTTGTC CTTCCAAGTC CGCCAGTTTA AGCGCCTTTA TGAACATATT AAAAATGACA 2700

AGTACCTTGK GGGTCAASGA CTCGGTGAAC TACGAACGGA AATCTGGGCA AACAAGGCTC 2760

ATCTCCACCA CCTCCACAGT CATCCCAAGA ATAAAGTAGT TGTCTCAACA ACTTGACCTT 2820

CCCCTTTACA TGTCCTTTTT TGTGGACTTC TCTCTTKGGA GATTTTTCCC AGTGATCTCT 2880

CAGCGTKGTT TTTAAGTTAA AKGTATTKGA CTTGTGTTCT CAGCATTCAG AGAGCAGCGG 2940

TGTAAGATTC TGCTGTTCTC CCTGGATCTT CTGACATKAC TGCTGTCTGA GATTTGTATA 3000

TGKGTAAATA CAAGTTCCTT GATACCCTAA AACCTTGGAT TAAACAGAAT GTGCATKGTA 3060

CATCTTTAAA CAAAATGKAT ATTAATTTAT TAAAAAAAAA AAAAAAAAA 3109

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 750 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Met Gly Glu His Gln Phe Gly Trp Ser Met Leu Pro His Arg Ser Met
1 5 10 15

Leu Arg Gly Ile Thr Lys Met Arg Leu Gln Gln Glu Glu Met Val Gln
20 25 30

Lys Met Leu Leu Leu Ile Ser Leu Leu Thr His Gln Leu Arg Thr Gln
35 40 45

Trp Trp Gly Lys Thr Leu Met Pro Arg Met Thr Arg Gln Lys Arg Arg
50 55 60

Arg Arg Thr Met Arg Arg Lys Met Thr Leu Val Trp Arg Met Ala Ala
65 70 75 80

Asp Ala Asn Asn Gly Ala Gln Asp Asp Met Asn Trp Asn Ala Leu Glu
85 90 95

Trp Asp Arg Ala Ala Glu Glu Leu Thr Trp Glu Arg Met Leu Gly Leu
100 105 110

Asp Gly Ser Leu Val Phe Leu Glu His Val Phe Trp Val Val Ser Leu
115 120 125

Asn Thr Leu Phe Ile Leu Val Phe Ala Phe Cys Pro Tyr His Ile Gly
130 135 140

His Phe Ser Leu Val Gly Leu Gly Phe Glu Glu His Val Gln Ala Ser
145 150 155 160

His Phe Glu Gly Leu Ile Thr Thr Ile Val Gly Tyr Ile Leu Leu Ala
165 170 175

Ile Thr Leu Ile Ile Cys His Gly Leu Ala Thr Leu Val Lys Phe His
180 185 190

Arg Ser Arg Arg Leu Leu Gly Val Cys Tyr Ile Val Val Lys Val Ser
195 200 205

Leu Leu Val Val Val Glu Ile Gly Val Phe Pro Leu Ile Cys Gly Trp
210 215 220

Trp Leu Asp Ile Cys Ser Leu Glu Met Phe Asp Ala Thr Leu Lys Asp
225 230 235 240

Arg Glu Leu Ser Phe Gln Ser Ala Pro Gly Thr Thr Met Phe Leu His
245 250 255

Trp Leu Val Gly Met Val Tyr Val Phe Tyr Phe Ala Ser Phe Ile Leu
260 265 270

Leu Leu Arg Glu Val Leu Arg Pro Gly Val Leu Trp Phe Leu Arg Asn
275 280 285

Leu Asn Asp Pro Asp Phe Asn Pro Val Gln Glu Met Ile His Leu Pro
290 295 300

Ile Tyr Arg His Leu Arg Arg Phe Ile Leu Ser Val Ile Val Phe Gly
305 310 315 320

Ser Ile Val Leu Leu Met Leu Trp Leu Pro Ile Arg Ile Ile Lys Ser
325 330 335

Val Leu Pro Asn Phe Leu Pro Tyr Asn Val Met Leu Tyr Ser Asp Ala
340 345 350

Pro Val Ser Glu Leu Ser Leu Glu Leu Leu Leu Leu Gln Val Val Leu
355 360 365

Pro Ala Leu Leu Glu Gln Gly His Thr Arg Gln Trp Leu Lys Gly Leu
370 375 380

Val Arg Ala Trp Thr Val Thr Ala Gly Tyr Leu Leu Asp Leu His Ser
385 390 395 400

Tyr Leu Leu Gly Asp Gln Glu Glu Asn Glu Asn Ser Ala Asn Gln Gln
405 410 415

Val Asn Asn Asn Gln His Ala Arg Asn Asn Asn Ala Ile Pro Val Val

420 425 430

Gly Glu Gly Leu His Ala Ala His Gln Ala Ile Leu Gln Gln Gly Gly
435 440 445

Pro Val Gly Phe Gln Pro Tyr Arg Arg Pro Leu Asn Phe Pro Leu Arg
450 455 460

Ile Phe Leu Leu Ile Val Phe Met Cys Ile Thr Leu Leu Ile Ala Ser
465 470 475 480

Leu Ile Cys Leu Thr Leu Pro Val Phe Ala Gly Arg Trp Leu Met Ser
485 490 495

Phe Trp Thr Gly Thr Ala Lys Ile His Glu Leu Tyr Thr Ala Ala Cys
500 505 510

Gly Leu Tyr Val Cys Trp Leu Thr Ile Arg Ala Val Thr Val Met Val
515 520 525

Ala Trp Met Pro Gln Gly Arg Arg Val Ile Phe Gln Lys Val Lys Glu
530 535 540

Trp Ser Leu Met Ile Met Lys Thr Leu Ile Val Ala Val Leu Leu Ala
545 550 555 560

Gly Val Val Pro Leu Leu Leu Gly Leu Leu Phe Glu Leu Val Ile Val
565 570 575

Ala Pro Leu Arg Val Pro Leu Asp Gln Thr Pro Leu Phe Tyr Pro Trp
580 585 590

Gln Asp Trp Ala Leu Gly Val Leu His Ala Lys Ile Ile Ala Ala Ile
595 600 605

Thr Leu Met Gly Pro Gln Trp Trp Leu Lys Thr Val Ile Glu Gln Val
610 615 620

Tyr Ala Asn Gly Ile Arg Asn Ile Asp Leu His Tyr Ile Val Arg Lys
625 630 635 640

Leu Ala Ala Pro Val Ile Ser Val Leu Leu Leu Ser Leu Cys Val Pro
645 650 655

Tyr Val Ile Ala Ser Gly Val Val Pro Leu Leu Gly Val Thr Ala Glu
660 665 670

Met Gln Asn Leu Val His Arg Arg Ile Tyr Pro Phe Leu Leu Met Val
675 680 685

Val Val Leu Met Ala Ile Leu Ser Phe Gln Val Arg Gln Phe Lys Arg
690 695 700

Leu Tyr Glu His Ile Lys Asn Asp Lys Tyr Leu Xaa Gly Gln Xaa Leu
705 710 715 720

Gly Glu Leu Arg Thr Glu Ile Trp Ala Asn Lys Ala His Leu His His
725 730 735

Leu His Ser His Pro Lys Asn Lys Val Val Val Ser Thr Thr
740 745 750

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

TNTTTGAAGT TTCTCCCTCT CATTCTGAG 29

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GNTTCTCCAC GTAGTTGGTT TTCCTCAGT 29

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CNACATGACG TGAGCTGGTG ATCCATGAA 29

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

ANTTGGGCTC TGCCGTCCAG AAAGGTTTG 29

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GNAGCTACGC GGACTTGCAG AGGTTTTAT 29

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

TNGGTGAGAG AATAGAGACT GGCTGGGAA 29

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

ANGAGCCGAC CAGACGTGCA AGATGCATA 29

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

ANCTGACCAG GACTCTCCTT CCACAGAGG 29

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

TNTAGGCGGA AATGGTGCTG GCCATGGTG 29

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

ANATATCCAG CCACCAACCA CAAATGAGA 29

What is claimed is:

1. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 185 to nucleotide 1600;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 1403 to nucleotide 1600;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 1 to nucleotide 850;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone do15_4 deposited under accession number ATCC 98468;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone do15_4 deposited under accession number ATCC 98468;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone do15_4 deposited under accession number ATCC 98468;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone do15_4 deposited under accession number ATCC 98468;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:2;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

2. The polynucleotide of claim 1 wherein said polynucleotide is operably linked to at least one expression control sequence.

3. A host cell transformed with the polynucleotide of claim 2.

4. The host cell of claim 3, wherein said cell is a mammalian cell.

5. A process for producing a protein encoded by the polynucleotide of claim 2, which process comprises:

(a) growing a culture of the host cell of claim 3 in a suitable culture medium; and

(b) purifying said protein from the culture.

6. A protein produced according to the process of claim 5.

7. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:2;

(b) the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 222;

(c) fragments of the amino acid sequence of SEQ ID NO:2 comprising eight consecutive amino acids of SEQ ID NO:2; and

(d) the amino acid sequence encoded by the cDNA insert of clone do15_4 deposited under accession number ATCC 98468;

the protein being substantially free from other mammalian proteins.

8. The protein of claim 7, wherein said protein comprises the amino acid sequence of SEQ ID NO:2.

9. The protein of claim 7, wherein said protein comprises the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 222.

10. A composition comprising the protein of claim 7 and a pharmaceutically acceptable carrier.

11. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:1.

12. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 47 to nucleotide 2065;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 1086 to nucleotide 1848;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone dx290_1 deposited under accession number ATCC 98468;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone dx290_1 deposited under accession number ATCC 98468;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone dx290_1 deposited under accession number ATCC 98468;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone dx290_1 deposited under accession number ATCC 98468;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:4;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

13. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:4;

(b) the amino acid sequence of SEQ ID NO:4 from amino acid 312 to amino acid 600;

(c) fragments of the amino acid sequence of SEQ ID NO:4 comprising eight consecutive amino acids of SEQ ID NO:4; and

(d) the amino acid sequence encoded by the cDNA insert of clone dx290_1 deposited under accession number ATCC 98468;

the protein being substantially free from other mammalian proteins.

14. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:3.

15. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 107 to nucleotide 724;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 218 to nucleotide 724;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 536 to nucleotide 866;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ek390_4 deposited under accession number ATCC 98468;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ek390_4 deposited under accession number ATCC 98468;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone ek390_4 deposited under accession number ATCC 98468;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone ek390_4 deposited under accession number ATCC 98468;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:6;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

16. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:6;

(b) the amino acid sequence of SEQ ID NO:6 from amino acid 6 to amino acid 92;

(c) fragments of the amino acid sequence of SEQ ID NO:6 comprising eight consecutive amino acids of SEQ ID NO:6; and

(d) the amino acid sequence encoded by the cDNA insert of clone ek390_4 deposited under accession number ATCC 98468;

the protein being substantially free from other mammalian proteins.

17. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:5.

18. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 31 to nucleotide 1230;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 289 to nucleotide 1230;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 344 to nucleotide 1119;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone er471_7 deposited under accession number ATCC 98468;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone er471_7 deposited under accession number ATCC 98468;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone er471_7 deposited under accession number ATCC 98468;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone er471_7 deposited under accession number ATCC 98468;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:8;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

19. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:8;

(b) the amino acid sequence of SEQ ID NO:8 from amino acid 111 to amino acid 363;

(c) fragments of the amino acid sequence of SEQ ID NO:8 comprising eight consecutive amino acids of SEQ ID NO:8; and

(d) the amino acid sequence encoded by the cDNA insert of clone er471_7 deposited under accession number ATCC 98468;

the protein being substantially free from other mammalian proteins.

20. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:7.

21. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 62 to nucleotide 322;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 571 to nucleotide 878;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone fs40_3 deposited under accession number ATCC 98468;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone fs40_3 deposited under accession number ATCC 98468;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone fs40_3 deposited under accession number ATCC 98468;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone fs40_3 deposited under accession number ATCC 98468;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:10;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:10;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

22. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:10;

(b) fragments of the amino acid sequence of SEQ ID NO:10 comprising eight consecutive amino acids of SEQ ID NO:10; and

(c) the amino acid sequence encoded by the cDNA insert of clone fs40_3 deposited under accession number ATCC 98468;

the protein being substantially free from other mammalian proteins.

23. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:9.

24. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 43 to nucleotide 1671;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 112 to nucleotide 1671;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 224 to nucleotide 679;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ga63_6 deposited under accession number ATCC 98468;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ga63_6 deposited under accession number ATCC 98468;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone ga63_6 deposited under accession number ATCC 98468;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone ga63_6 deposited under accession number ATCC 98468;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:12;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

25. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:12;

(b) the amino acid sequence of SEQ ID NO:12 from amino acid 62 to amino acid 212;

(c) fragments of the amino acid sequence of SEQ ID NO:12 comprising eight consecutive amino acids of SEQ ID NO:12; and

(d) the amino acid sequence encoded by the cDNA insert of clone ga63_6 deposited under accession number ATCC 98468;

the protein being substantially free from other mammalian proteins.

26. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:11.

27. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 17 to nucleotide 523;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 77 to nucleotide 523;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 1 to nucleotide 392;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone gm335_4 deposited under accession number ATCC 98468;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone gm335_4 deposited under accession number ATCC 98468;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone gm335_4 deposited under accession number ATCC 98468;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone gm335_4 deposited under accession number ATCC 98468;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:14;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:14;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

28. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:14;

(b) the amino acid sequence of SEQ ID NO:14 from amino acid 1 to amino acid 125;

(c) fragments of the amino acid sequence of SEQ ID NO:14 comprising eight consecutive amino acids of SEQ ID NO:14; and

(d) the amino acid sequence encoded by the cDNA insert of clone gm335_4 deposited under accession number ATCC 98468;

the protein being substantially free from other mammalian proteins.

29. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:13.

30. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 2 to nucleotide 991;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 62 to nucleotide 991;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 2 to nucleotide 504;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone hy370_9 deposited under accession number ATCC 98468;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone hy370_9 deposited under accession number ATCC 98468;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone hy370_9 deposited under accession number ATCC 98468;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone hy370_9 deposited under accession number ATCC 98468;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:16;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

31. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:16;

(b) the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 167;

(c) fragments of the amino acid sequence of SEQ ID NO:16 comprising eight consecutive amino acids of SEQ ID NO:16; and

(d) the amino acid sequence encoded by the cDNA insert of clone hy370_9 deposited under accession number ATCC 98468;

the protein being substantially free from other mammalian proteins.

32. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:15.

33. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 77 to nucleotide 616;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 164 to nucleotide 616;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 1 to nucleotide 415;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ie47_4 deposited under accession number ATCC 98468;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ie47_4 deposited under accession number ATCC 98468;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone ie47_4 deposited under accession number ATCC 98468;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone ie47_4 deposited under accession number ATCC 98468;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:18;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:18;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

34. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:18;

(b) the amino acid sequence of SEQ ID NO:18 from amino acid 1 to amino acid 113;

(c) fragments of the amino acid sequence of SEQ ID NO:18 comprising eight consecutive amino acids of SEQ ID NO:18; and

(d) the amino acid sequence encoded by the cDNA insert of clone ie47_4 deposited under accession number ATCC 98468;

the protein being substantially free from other mammalian proteins.

35. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:17.

36. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 564 to nucleotide 2813;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 705 to nucleotide 2813;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 793 to nucleotide 1628;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone s195_10 deposited under accession number ATCC 98468;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone s195_10 deposited under accession number ATCC 98468;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone s195_10 deposited under accession number ATCC 98468;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone s195_10 deposited under accession number ATCC 98468;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:20;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20 having biological activity, the fragment comprising eight consecutive amino acids of SEQ ID NO:20;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

37. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:20;

(b) the amino acid sequence of SEQ ID NO:20 from amino acid 78 to amino acid 355;

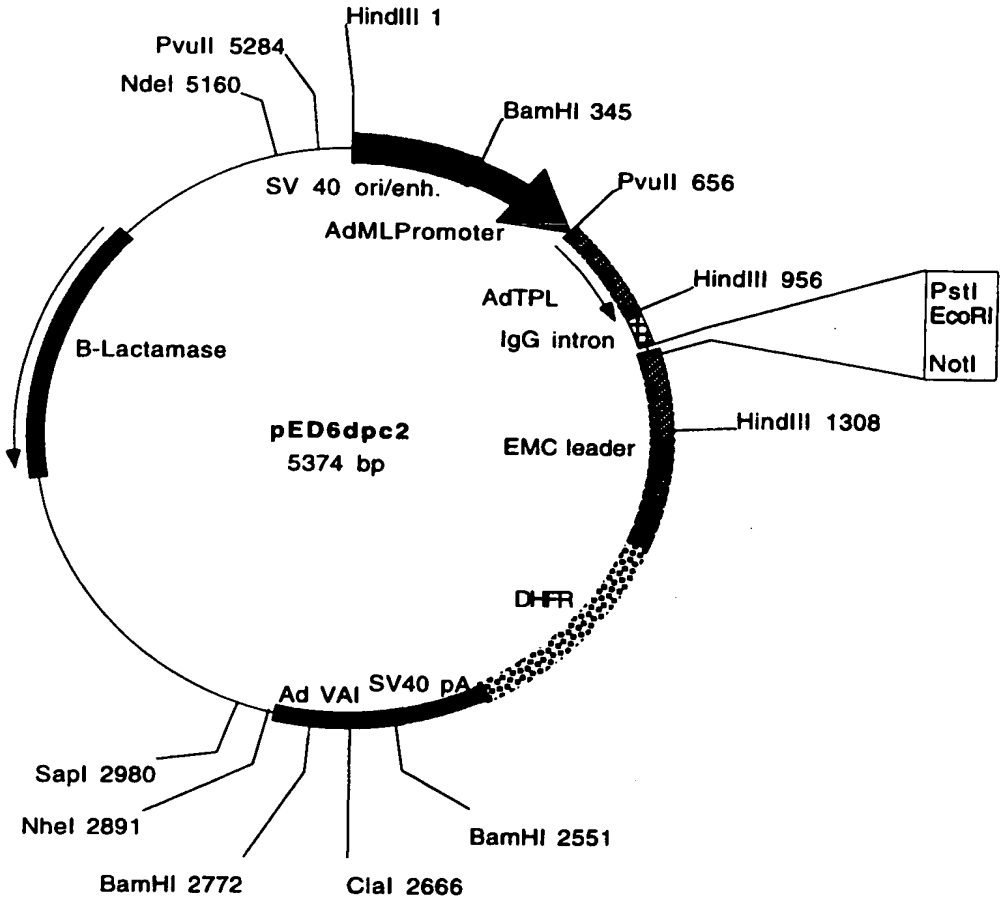
(c) fragments of the amino acid sequence of SEQ ID NO:20 comprising eight consecutive amino acids of SEQ ID NO:20; and

(d) the amino acid sequence encoded by the cDNA insert of clone s195_10 deposited under accession number ATCC 98468;

the protein being substantially free from other mammalian proteins.

38. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:19.

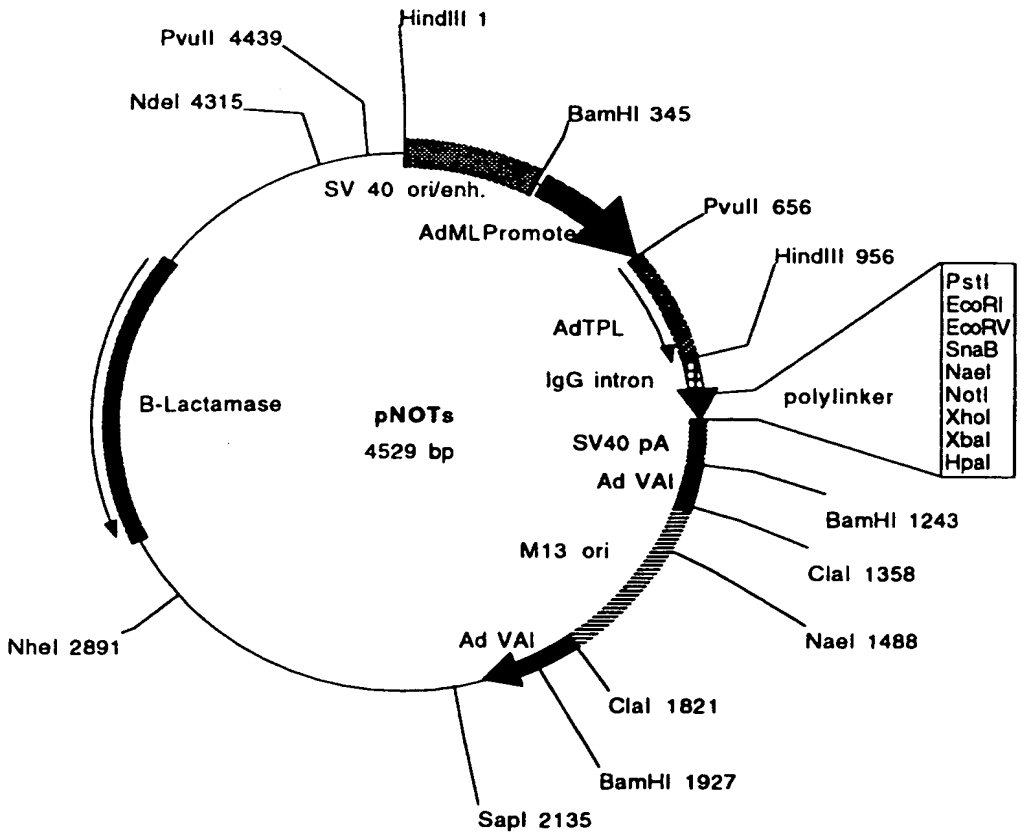
FIGURE 1A



Plasmid name: pED6dpc2
Plasmid size: 5374 bp

Comments/References: pED6dpc2 is derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al.(1991), NAR 19: 4485-4490.

FIGURE 1B



Plasmid name: pNOTs
Plasmid size: 4529 bp

Comments/References: pNOTs is a derivative of pMT2 (Kaufman et al,1989. Mol.Cell.Biol.9:1741-1750). DHFR was deleted and a new polylinker was inserted between EcoRI and HpaI. M13 origin of replication was ins rted in the ClaI site. SST cDNAs are cloned b tw en EcoRI and NotI

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/12516

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :C 07K 1/OO; CO7H 21/02; A61K 39/00; C12N 1/20
US CL : 530/350; 536/23.1; 424/184.1; 435/325, 252.3

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 530/350; 536/23.1; 424/184.1; 435/325, 252.3

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS
search terms: secreted proteins, DNA, proteins, expression vectors

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,536,637 A (JACOBS) 16 June 1996, see entire dicument.	1-38
Y,E	US 5,770,209 A (GROTENDORST et al) 23 June 1998, see entire document.	1-38
Y	US 4,798,885 A (MASON et al) 17 January 1989, see entire document.	1-38
Y,E	US 5,773,586 A (GOSPODAROWICZ et al) 30 June 1998, see entire document.	1-38
Y,E	US 5,712,116 A (JACOBS) 27 June 1998, see entire document.	1-38
X,P	US 5,708,157 A (JACOBS et al) 13 January 1998, see entire document.	1-38

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"B"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 02 SEPTEMBER 1998	Date of mailing of the international search report 16 OCT 1998
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer Lynette F. Smith Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/12516

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,E	US 5,792,628 A (BOWMAN) 11 August 1998, see entire document.	1-38
X,P	US 5,707,829 A (JACOBS et al) 13 January 1998, see entire document.	1-38